



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Harrington, Lea A.
Robinson, Murray O.
- (ii) TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS
- (iii) NUMBER OF SEQUENCES: 44
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Amgen Inc.
 - (B) STREET: One Amgen Center Drive
 - (C) CITY: Thousand Oaks
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 91320-1789
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/951,733
 - (B) FILING DATE: 16-OCT-1997
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/873,039
 - (B) FILING DATE: 11-JUN-1997
- (viii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/751,189
 - (B) FILING DATE: 15-NOV-1996
- (ix) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Oleski, Nancy A.
 - (B) REGISTRATION NUMBER: 34,688
 - (C) REFERENCE/DOCKET NUMBER: A-433B
- (x) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (805) 447-6504
 - (B) TELEFAX: (805) 499-8011

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7881 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGAAAAAC TCCATGGGCA TGTGTCTGCC CATCCAGACA TCCTCTCCTT GGAGAACCGG

60

TGCCTGGCTA	TGCTCCCTGA	CTTACAGCCC	TTGGAGAAC	TACATCAGCA	TGTATCTACC	120
CACTCAGATA	TCCTCTCCTT	GAAGAACCGAG	TGCCTAGCCA	CGCTTCCTGA	CCTGAAGACC	180
ATGGAAAAAC	CACATGGATA	TGTGTCTGCC	CACCCAGACA	TCCTCTCCTT	GGAGAACCGAG	240
TGCCTGGCCA	CACTTCTGA	CCTGAAGACC	ATGGAGAAC	CACATGGACA	TGTTTCTGCC	300
CACCCAGACA	TCCTCTCCTT	GGAGAACCGG	TGCCTGGCCA	CCCTCCCTAG	TCTAAAGAGC	360
ACTGTGTCTG	CCAGCCCCTT	GTTCCAGAGT	CTACAGATAT	CTCACATGAC	GCAAGCTGAT	420
TTGTACCGTG	TGAACAAACAG	CAATTGCCTG	CTCTCTGAGC	CTCCAAGTTG	GAGGGCTCAG	480
CATTTCTCTA	AGGGACTAGA	CCTTTCAACC	TGCCCTATAG	CCCTGAAATC	CATCTCTGCC	540
ACAGAGACAG	CTCAGGAAGC	AACTTTGGGT	CGTTGGTTG	ATTCAAGAAGA	GAAGAAAGGG	600
GCAGAGACCC	AAATGCCTTC	TTATAGTCTG	AGCTTGGGAG	AGGAGGAGGA	GGTGGAGGAT	660
CTGGCCGTGA	AGCTCACCTC	TGGAGACTCT	GAATCTCATC	CAGAGCCTAC	TGACCATGTC	720
CTTCAGGAAA	AGAAGATGGC	TCTACTGAGC	TTGCTGTGCT	CTACTCTGGT	CTCAGAAGTA	780
AACATGAACA	ATACATCTGA	CCCCACCCTG	GCTGCCATT	TTGAAATCTG	TCGTGAACCT	840
GCCCTCCTGG	AGCCTGAGTT	TATCCTCAAG	GCATCTTGT	ATGCCAGGCA	GCAGCTGAAC	900
GTCCGGAATG	TGGCCAATAA	CATCTTGCC	ATTGCTGCTT	TCTTGCCGGC	GTGTCGCCCC	960
CACCTGCGAC	GATATTCTG	TGCCATTGTC	CAGCTGCCTT	CTGACTGGAT	CCAGGTGGCT	1020
GAGCTTTACC	AGAGCCTGGC	TGAGGGAGAT	AAGAATAAGC	TGGTCCCCCT	GCCCGCCTGT	1080
CTCCGTACTG	CCATGACGGA	CAAATTGCC	CAGTTGACG	AGTACCAGCT	GGCTAAGTAC	1140
AACCCTCGGA	AGCACCGGGC	CAAGAGACAC	CCCCGCCGGC	CACCCCGCTC	TCCAGGGATG	1200
GAGCCTCCAT	TTTCTCACAG	ATGTTTCCA	AGGTACATAG	GGTTTCTCAG	AGAAGAGCAG	1260
AGAAAGTTG	AGAAGGCCGG	TGATACAGTG	TCAGAGAAA	AGAACCTCC	AAGGTTCACCC	1320
CTGAAGAACG	TGGTCAGCG	ACTGCACATC	CACAAGCCTG	CCCAGCACGT	TCAAGCCCTG	1380
CTGGGTTACA	GATACCCCTC	CAACCTACAG	CTCTTTCTC	GAAGTCGCCT	TCCTGGGCCT	1440
TGGGATTCTA	GCAGAGCTGG	GAAGAGGATG	AAGCTGTCTA	GGCCAGAGAC	CTGGGAGCGG	1500
GAGCTGAGCC	TACGGGGAA	CAAAGCGTCG	GTCTGGGAGG	AACTCATTGA	AAATGGGAAG	1560
CTTCCCTTCA	TGGCCATGCT	TCGGAACCTG	TGCAACCTGC	TGCGGGTTGG	AATCAGTTCC	1620
CGCCACCATG	AGCTCATTCT	CCAGAGACTC	CAGCATGGGA	AGTCGGTGAT	CCACAGTCGG	1680
CAGTTCCAT	TCAGATTCT	TAACGCCAT	GATGCCATTG	ATGCCCTCGA	GGCTCAACTC	1740
AGAAATCAAG	CATTGCCCTT	TCCTTCGAAT	ATAACACTGA	TGAGGCCGGAT	ACTAACTAGA	1800
AATGAAAAGA	ACCGTCCCAG	CGGGAGGTTT	CTTGCCACC	TAAGCCGTCA	GCAGCTTCGT	1860
ATGGCAATGA	GGATACCTGT	GTTGTATGAG	CAGCTCAAGA	GGGAGAAGCT	GAGAGTACAC	1920

AAGGCCAGAC	AGTGGAAATA	TGATGGTGAG	ATGCTGAACA	GGTACCGACA	G GCCCTAGAG	1980
ACAGCTGTGA	ACCTCTCTGT	GAAGCACAGC	CTGCCCTGC	TGCCAGGCCG	CACTGTCTTG	2040
GTCTATCTGA	CAGATGCTAA	TGCAGACAGG	CTCTGTCCAA	AGAGCAACCC	ACAAGGGCCC	2100
CCGCTGAACT	ATGCACTGCT	GTTGATTGGG	ATGATGATCA	CGAGGGCGGA	GCAGGTGGAC	2160
GTCGTGCTGT	GTGGAGGTGA	CACTCTGAAG	ACTGCAGTGC	TTAAGGCAGA	AGAAGGCATC	2220
CTGAAGACTG	CCATCAAGCT	CCAGGCTCAA	GTCCAGGAGT	TTGATGAAAA	TGATGGATGG	2280
TCCCTGAATA	CTTTTGGAA	ATACCTGCTG	TCTCTGGCTG	GCCAAAGGGT	TCCTGTGGAC	2340
AGGGTCATCC	TCCTTGGCCA	AAGCATGGAT	GATGGAATGA	TAAATGTGGC	CAAACAGCTT	2400
TACTGGCAGC	GTGTGAATTC	CAAGTGCCTC	TTTGTGTTGA	TCCTCCTAACG	AAGGGTACAA	2460
TACCTGTCAA	CAGATTTGAA	TCCCATTGAT	GTGACACTCT	CAGGCTGTAC	TGATGCGATA	2520
CTGAAGTTCA	TTGCAGAGCA	TGGGGCCTCC	CATCTTCTGG	AACATGTGGG	CCAAATGGAC	2580
AAAATATTCA	AGATTCCACC	ACCCCCCAGGA	AAGACAGGGG	TCCAGTCTCT	CCGGCCACTG	2640
GAAGAGGACA	CTCCAAGCCC	CTTGGCTCCT	GTTTCCCAGC	AAGGATGGCG	CAGCATCCGG	2700
CTTTTCATT	CATCCACTTT	CCGAGACATG	CACGGGGAGC	GGGACCTGCT	GCTGAGGTCT	2760
GTGCTGCCAG	CACTGCAGGC	CCGAGCGGCC	CCTCACCGTA	TCAGCCTTCA	CGGAATCGAC	2820
CTCCGCTGGG	GC GTC ACTGA	GGAGGAGACC	CGTAGGAACA	GACAAC TGGA	AGTGTGCCTT	2880
GGGGAGGTGG	AGAACGCACA	GCTGTTGTG	GGGATTCTGG	GCTCCCGTTA	TGGATACATT	2940
CCCCCCAGCT	ACAACCTTCC	TGACCATCCA	CACTTCCACT	GGGCCAGCA	GTACCCTTCA	3000
GGCGCTCTG	TGACAGAGAT	GGAGGTGATG	CAGTTCCCTGA	ACCGGAACCA	ACGTCTGCAG	3060
CCCTCTGCC	AAGCTCTCAT	CTACTTCCGG	GATTCCAGCT	TCCTCAGCTC	TGTGCCAGAT	3120
GCCTGGAAAT	CTGACTTTGT	TTCTGAGTCT	GAAGAGGCCG	CATGTCGGAT	CTCAGAACTG	3180
AAGAGCTACC	TAAGCAGACA	GAAAGGGATA	ACCTGCCGCA	GATA CCCCTG	TGAGTGGGG	3240
GGTGTGGCAG	CTGGCCGGCC	CTATGTTGGC	GGGCTGGAGG	AGTTTGGCA	GTTGGTTCTG	3300
CAGGATGTAT	GGAAATATGAT	CCAGAAGCTC	TACCTGCAGC	CTGGGCCCT	GCTGGAGCAG	3360
CCAGTGTCCA	TCCCAGACGA	TGACTTGGTC	CAGGCCACCT	TCCAGCAGCT	GCAGAAGCCA	3420
CCGAGTCCTG	CCCAGGCCACG	CCTTCTTCAG	GACACAGTGC	AACAGCTGAT	GCTGCCCTCAC	3480
GGAAGGCTGA	GCCTGGTGAC	GGGGCAGTCA	GGACAGGGCA	AGACAGCCTT	CCTGGCATCT	3540
CTTGTGTCAG	CCCTGCAGGC	TCCTGATGGG	GCCAAAGGTGG	CACCATTAGT	CTTCTTCCAC	3600
TTTCTGGGG	CTCGTCCCTGA	CCAGGGTCTT	GCCCTCACTC	TGCTCAGACG	CCTCTGTACC	3660
TATCTGCGTG	GCCA ACT AAA	AGAGCCAGGT	GCCCTCCCCA	GCACCTACCG	AAGCCTGGTG	3720
TGGGAGCTGC	AGCAGAGGCT	GCTGCCAAG	TCTGCTGAGT	CCCTGCATCC	TGGCCAGACC	3780
CAGGT CCTGA	TCATCGATGG	GGCTGATAGG	TTAGTGGACC	AGAATGGCA	GCTGATTTC	3840

GAATGAGCG	3900
GCAGGGCTAG	3960
CTGGAGGCCT	4020
CTGGAGGAGT	4080
CGGCCGCTCT	4140
GTGTCTGAGA	4200
AGCACACTGG	4260
GTCACACGGA	4320
CTACCGAAGG	4380
TACCCCATGG	4440
CCTCTGGAGC	4500
AAACGTTGCT	4560
CAGCTCTGGA	4620
TTCCATTACCA	4680
TTGGAGGCC	4740
GCTGACGTTG	4800
CCCCGGCTCC	4860
GCCTCGCTGC	4920
CGGACCATGA	4980
GCTGTGGCCT	5040
TACCTGTTGG	5100
ATCTCTGCTT	5160
CTGGAGCTCT	5220
CAAATCACTG	5280
GGATGCCCTAA	5340
AAGTCCCTGA	5400
GCTGGCAGCA	5460
CCCGGAGCCT	5520
CGGCTGGACA	5580
CCTGCCACC	5640
ATGGCTTGT	5700
TGCTGCTGCG	
CTTTCCCTGC	
ATGCGGGTTG	
CCAGTTACTG	
GTATGGTGG	
GCTGTGGCC	
TGGCGAGAAG	
GGGCACGGCT	
GGCTGCCTTC	
ATGGCCTTGT	
CGCTGCTGAG	
CTCCCTGATG	
GGCCCCTGAG	
AACAGCAGCT	
AAACGTTGCT	
ATGGGAAGAG	
GCCAGGGCTA	
GAGGACACGG	
CACACATCCT	
CATTGCAGCT	
CGCTGATGCC	
TCAGGCACCT	
TCCGAAGTTG	
CCCTCCTGAG	
ACCTGCCTTA	
CCACCTGCTC	
CAGAGCGGGA	
ACCGTGGACT	
TCTTTCGAAG	
ACCTCCATGT	
GGTGGCTGCA	
CACTTGAAT	
TGGGTCTGGT	
CTCTCGGCTC	
TGCTTCTTCA	
GTCCCCAAAG	
AGGAACAAAA	
GCTCCCCGAG	
CAGTGTTCG	
CACCTTCCTG	
AGGCAGCAGG	
CTTCAATCCT	
CAGCCAGTAC	
GGCAGCCAAC	
CAGCCCCTGG	
ACTCACCTCT	
TTGCCACCAA	
TCTCCGGAG	
ATGGCACCTC	
CAACACACAC	
TACGATGGCT	
TAATAAACCC	
AAAATCAGCA	
AAGCTCCAGC	
CTGTCTCTGG	
CAGTTCCCTC	
ATCCCCTACT	
GCTGTGGCCT	
TCTCCACCAA	
TGGGCAAAGA	
GCAGCTGTGG	
GCACTGCCAA	
TGGGACAGTT	
TTGGCAGGAG	
GAGAAGTCTG	
TGGTGAGTGG	
CTGTGATGGA	
ACACTCTTC	
TTACTGCCTT	
CGACGGGCTC	
GGACCTGCA	
GCATGGTTGT	
CGGGTGCTGC	
AGACTAAGGC	
TCACCACTGAC	
GCTGCTGCCT	
GAGCCCAGAC	
TGCCGGCTGC	
TAGCCACCGT	
GTGCTTGGGA	
AGCTGTGGGA	
CACAGTCCGT	
GGGCAGCTGG	
CCTTCCAGCA	
CACCTACCCC	
ACTGTGTTGC	
CTTCCACCCA	
GAGGGGCAGG	
TAATAGCCAC	
AGGCAGCTGG	
TCAGGTTCTT	
CCAGGTGGAT	
GGGCTCAAAG	
TCACCAAGGA	
CCTGGGGGCA	
CTATCCGTAC	
CTTGGCCTTC	
AATGTGCCTG	
GGGGGGTTGT	
GGCTGTGGGC	
GTATGGTGG	
GCTGTGGCC	
TGGCGAGAAG	
GGGCACGGCT	
GGCTGCCTTC	
ATGGCCTTGT	
CGCTGCTGAG	
CTCCCTGATG	
GGCCCCTGAG	
AACAGCAGCT	
AAACGTTGCT	
ATGGGAAGAG	
GCCAGGGCTA	
GAGGACACGG	
CACACATCCT	
CATTGCAGCT	
CGCTGATGCC	
TCAGGCACCT	
TCCGAAGTTG	
CCCTCCTGAG	
ACCTGCCTTA	
CCACCTGCTC	
CAGAGCGGGA	
ACCGTGGACT	
TCTTTCGAAG	
ACCTCCATGT	
GGTGGCTGCA	
CACTTGAAT	
TGGGTCTGGT	
CTCTCGGCTC	
TGCTTCTTCA	
GTCCCCAAAG	
AGGAACAAAA	
GCTCCCCGAG	
CAGTGTTCG	
CACCTTCCTG	
AGGCAGCAGG	
CTTCAATCCT	
CAGCCAGTAC	
GGCAGCCAAC	
CAGCCCCTGG	
ACTCACCTCT	
TTGCCACCAA	
TCTCCGGAG	
ATGGCACCTC	
CAACACACAC	
TACGATGGCT	
TAATAAACCC	
AAAATCAGCA	
AAGCTCCAGC	
CTGTCTCTGG	
CAGTTCCCTC	
ATCCCCTACT	
GCTGTGGCCT	
TCTCCACCAA	
TGGGCAAAGA	
GCAGCTGTGG	
GCACTGCCAA	
TGGGACAGTT	
TTGGCAGGAG	
GAGAAGTCTG	
TGGTGAGTGG	
CTGTGATGGA	
ACACTCTTC	
TTACTGCCTT	
CGACGGGCTC	
GGACCTGCA	
GCATGGTTGT	
CGGGTGCTGC	
AGACTAAGGC	
TCACCACTGAC	
GCTGCTGCCT	
GAGCCCAGAC	
TGCCGGCTGC	
TAGCCACCGT	
GTGCTTGGGA	
AGCTGTGGGA	
CACAGTCCGT	
GGGCAGCTGG	
CCTTCCAGCA	
CACCTACCCC	
ACTGTGTTGC	
CTTCCACCCA	
GAGGGGCAGG	
TAATAGCCAC	
AGGCAGCTGG	
TCAGGTTCTT	
CCAGGTGGAT	
GGGCTCAAAG	
TCACCAAGGA	
CCTGGGGGCA	
CTATCCGTAC	
CTTGGCCTTC	
AATGTGCCTG	
GGGGGGTTGT	
GGCTGTGGGC	
GTATGGTGG	
GCTGTGGCC	
TGGCGAGAAG	
GGGCACGGCT	
GGCTGCCTTC	
ATGGCCTTGT	
CGCTGCTGAG	
CTCCCTGATG	
GGCCCCTGAG	
AACAGCAGCT	
AAACGTTGCT	
ATGGGAAGAG	
GCCAGGGCTA	
GAGGACACGG	
CACACATCCT	
CATTGCAGCT	
CGCTGATGCC	
TCAGGCACCT	
TCCGAAGTTG	
CCCTCCTGAG	
ACCTGCCTTA	
CCACCTGCTC	
CAGAGCGGGA	
ACCGTGGACT	
TCTTTCGAAG	
ACCTCCATGT	
GGTGGCTGCA	
CACTTGAAT	
TGGGTCTGGT	
CTCTCGGCTC	
TGCTTCTTCA	
GTCCCCAAAG	
AGGAACAAAA	
GCTCCCCGAG	
CAGTGTTCG	
CACCTTCCTG	
AGGCAGCAGG	
CTTCAATCCT	
CAGCCAGTAC	
GGCAGCCAAC	
CAGCCCCTGG	
ACTCACCTCT	
TTGCCACCAA	
TCTCCGGAG	
ATGGCACCTC	
CAACACACAC	
TACGATGGCT	
TAATAAACCC	
AAAATCAGCA	
AAGCTCCAGC	
CTGTCTCTGG	
CAGTTCCCTC	
ATCCCCTACT	
GCTGTGGCCT	
TCTCCACCAA	
TGGGCAAAGA	
GCAGCTGTGG	
GCACTGCCAA	
TGGGACAGTT	
TTGGCAGGAG	
GAGAAGTCTG	
TGGTGAGTGG	
CTGTGATGGA	
ACACTCTTC	
TTACTGCCTT	
CGACGGGCTC	
GGACCTGCA	
GCATGGTTGT	
CGGGTGCTGC	
AGACTAAGGC	
TCACCACTGAC	
GCTGCTGCCT	
GAGCCCAGAC	
TGCCGGCTGC	
TAGCCACCGT	
GTGCTTGGGA	
AGCTGTGGGA	
CACAGTCCGT	
GGGCAGCTGG	
CCTTCCAGCA	
CACCTACCCC	
ACTGTGTTGC	
CTTCCACCCA	
GAGGGGCAGG	
TAATAGCCAC	
AGGCAGCTGG	
TCAGGTTCTT	
CCAGGTGGAT	
GGGCTCAAAG	
TCACCAAGGA	
CCTGGGGGCA	
CTATCCGTAC	
CTTGGCCTTC	
AATGTGCCTG	
GGGGGGTTGT	
GGCTGTGGGC	
GTATGGTGG	
GCTGTGGCC	
TGGCGAGAAG	
GGGCACGGCT	
GGCTGCCTTC	
ATGGCCTTGT	
CGCTGCTGAG	
CTCCCTGATG	
GGCCCCTGAG	
AACAGCAGCT	
AAACGTTGCT	
ATGGGAAGAG	
GCCAGGGCTA	
GAGGACACGG	
CACACATCCT	
CATTGCAGCT	
CGCTGATGCC	
TCAGGCACCT	
TCCGAAGTTG	
CCCTCCTGAG	
ACCTGCCTTA	
CCACCTGCTC	
CAGAGCGGGA	
ACCGTGGACT	
TCTTTCGAAG	
ACCTCCATGT	
GGTGGCTGCA	
CACTTGAAT	
TGGGTCTGGT	
CTCTCGGCTC	
TGCTTCTTCA	
GTCCCCAAAG	
AGGAACAAAA	
GCTCCCCGAG	
CAGTGTTCG	
CACCTTCCTG	
AGGCAGCAGG	
CTTCAATCCT	
CAGCCAGTAC	
GGCAGCCAAC	
CAGCCCCTGG	
ACTCACCTCT	
TTGCCACCAA	
TCTCCGGAG	
ATGGCACCTC	
CAACACACAC	
TACGATGGCT	
TAATAAACCC	
AAAATCAGCA	
AAGCTCCAGC	
CTGTCTCTGG	
CAGTTCCCTC	
ATCCCCTACT	
GCTGTGGCCT	
TCTCCACCAA	
TGGGCAAAGA	
GCAGCTGTGG	
GCACTGCCAA	
TGGGACAGTT	
TTGGCAGGAG	
GAGAAGTCTG	
TGGTGAGTGG	
CTGTGATGGA	
ACACTCTTC	
TTACTGCCTT	
CGACGGGCTC	
GGACCTGCA	
GCATGGTTGT	
CGGGTGCTGC	
AGACTAAGGC	
TCACCACTGAC	
GCTGCTGCCT	
GAGCCCAGAC	
TGCCGGCTGC	
TAGCCACCGT	
GTGCTTGGGA	
AGCTGTGGGA	
CACAGTCCGT	
GGGCAGCTGG	
CCTTCCAGCA	
CACCTACCCC	
ACTGTGTTGC	
CTTCCACCCA	
GAGGGGCAGG	
TAATAGCCAC	
AGGCAGCTGG	
TCAGGTTCTT	
CCAGGTGGAT	
GGGCTCAAAG	
TCACCAAGGA	
CCTGGGGGCA	
CTATCCGTAC	
CTTGGCCTTC	
AATGTGCCTG	
GGGGGGTTGT	
GGCTGTGGGC	
GTATGGTGG	
GCTGTGGCC	
TGGCGAGAAG	
GGGCACGGCT	
GGCTGCCTTC	
ATGGCCTTGT	
CGCTGCTGAG	
CTCCCTGATG	
GGCCCCTGAG	
AACAGCAGCT	
AAACGTTGCT	
ATGGGAAGAG	
GCCAGGGCTA	
GAGGACACGG	
CACACATCCT	
CATTGCAGCT	
CGCTGATGCC	
TCAGGCACCT	
TCCGAAGTTG	
CCCTCCTGAG	
ACCTGCCTTA	
CCACCTGCTC	
CAGAGCGGGA	
ACCGTGGACT	
TCTTTCGAAG	
ACCTCCATGT	
GGTGGCTGCA	
CACTTGAAT	
TGGGTCTGGT	
CTCTCGGCTC	
TGCTTCTTCA	
GTCCCCAAAG	
AGGAACAAAA	
GCTCCCCGAG	
CAGTGTTCG	
CACCTTCCTG	
AGGCAGCAGG	
CTTCAATCCT	
CAGCCAGTAC	
GGCAGCCAAC	
CAGCCCCTGG	
ACTCACCTCT	
TTGCCACCAA	
TCTCCGGAG	
ATGGCACCTC	
CAACACACAC	
TACGATGGCT	
TAATAAACCC	
AAAATCAGCA	
AAGCTCCAGC	
CTGTCTCTGG	
CAGTTCCCTC	
ATCCCCTACT	
GCTGTGGCCT	
TCTCCACCAA	
TGGGCAAAGA	
GCAGCTGTGG	
GCACTGCCAA	
TGGGACAGTT	
TTGGCAGGAG	
GAGAAGTCTG	
TGGTGAGTGG	
CTGTGATGGA	
ACACTCTTC	
TTACTGCCTT	
CGACGGGCTC	
GGACCTGCA	
GCATGGTTGT	
CGGGTGCTGC	
AGACTAAGGC	
TCACCACTGAC	
GCTGCTGCCT	
GAGCCCAGAC	
TGCCGGCTGC	
TAGCCACCGT	
GTGCTTGGGA	
AGCTGTGGGA	
CACAGTCCGT	
GGGCAGCTGG	
CCTTCCAGCA	
CACCTACCCC	
ACTGTGTTGC	
CTTCCACCCA	
GAGGGGCAGG	
TAATAGCCAC	
AGGCAGCTGG	
TCAGGTTCTT	
CCAGGTGGAT	
GGGCTCAAAG	
TCACCAAGGA	
CCTGGGGGCA	
CTATCCGTAC	
CTTGGCCTTC	
AATGTGCCTG	
GGGGGGTTGT	
GGCTGTGGGC	
GTATGGTGG	
GCTGTGGCC	
TGGCGAGAAG	
GGGCACGGCT	
GGCTGCCTTC	
ATGGCCTTGT	
CGCTGCTGAG	
CTCCCTGATG	
GGCCCCTGAG	
AACAGCAGCT	
AAACGTTGCT	
ATGGGAAGAG	
GCCAGGGCTA	
GAGGACACGG	
CACACATCCT	
CATTGCAGCT	
CGCTGATGCC	
TCAGGCACCT	
TCCGAAGTTG	
CCCTCCTGAG	
ACCTGCCTTA	
CCACCTGCTC	
CAGAGCGGGA	
ACCGTGGACT	
TCTTTCGAAG	
ACCTCCATGT	
GGTGGCTGCA	
CACTTGAAT	
TGGGTCTGGT	
CTCTCGGCTC	
TGCTTCTTCA	
GTCCCCAAAG	
AGGAACAAAA	
GCTCCCCGAG	
CAGTGTTCG	
CACCTTCCTG	
AGGCAGCAGG	
CTTCAATCCT	
CAGCCAGTAC	
GGCAGCCAAC	
CAGCCCCTGG	
ACTCACCTCT	
TTGCCACCAA	
TCTCCGGAG	
ATGGCACCTC	
CAACACACAC	
TACGATGGCT	
TAATAAACCC	
AAAATCAGCA	
AAGCTCCAGC	
CTGTCTCTGG	
CAGTTCCCTC	
ATCCCCTACT	
GCTGTGGCCT	
TCTCCACCAA	
TGGGCAAAGA	
GCAGCTGTGG	
GCACTGCCAA	
TGGGACAGTT	
TTGGCAGGAG	
GAGAAGTCTG	
TGGTGAGTGG	
CTGTGATGGA	
ACACTCTTC	
TTACTGCCTT	
CGACGGGCTC	
GGACCTGCA	
GCATGGTTGT	
CGGGTGCTGC	
AGACTAAGGC	
TCACCACTGAC	
GCTGCTGCCT	
GAGCCCAGAC	
TGCCGGCTGC	
TAGCCACCGT	
GTGCTTGGGA	
AGCTGTGGGA	
CACAGTCCGT	
GGGCAGCTGG	
CCTTCCAGCA	
CACCTACCCC	
ACTGTGTTGC	
CTTCCACCCA	
GAGGGGCAGG	
TAATAGCCAC	
AGGCAGCTGG	
TCAGGTTCTT	
CCAGGTGGAT	
GGGCTCAAAG	
TCACCAAGGA	
CCTGGGGGCA	
CTATCCGTAC	
CTTGGCCTTC	
AATGTGCCTG	
GGGGGGTTGT	
GGCTGTGGGC	
GTATGGTGG	
GCTGTGGCC	
TGGCGAGAAG	
GGGCACGGCT	
GGCTGCCTTC	
ATGGCCTTGT	
CGCTGCTGAG	

ACGGCTGGAG	AGGATGGCAA	GGTCAGGTG	TGGTCAGGGT	CTCTGGTCG	GCCCCGTGGG	5760
CACCTGGTT	CCCTTCTCT	CTCTCCTGCC	CTCTCTGTGG	CACTCAGCCC	AGATGGTGAT	5820
CGGGTGGCTG	TTGGATATCG	AGCGGATGGC	ATTAGGATCT	ACAAAATCTC	TTCAGGTTCC	5880
CAGGGGGCTC	AGGGTCAGGC	ACTGGATGTG	GCAGTGTCCG	CCCTGGCCTG	GCTAAGCCCC	5940
AAGGTATTGG	TGAGTGGTGC	AGAAGATGGG	TCCTTGCAGG	GCTGGCACT	CAAGGAATGC	6000
TCCCTTCAGT	CCCTCTGGCT	CCTGTCCAGA	TTCCAGAAGC	CTGTGCTAGG	ACTGGCCACT	6060
TCCCAGGAGC	TCTTGGCTTC	TGCCTCAGAG	GATTTCACAG	TGCAGCTGTG	GCCAAGGCAG	6120
CTGCTGACGC	GGCCACACAA	GGCAGAACAC	TTTCCCTGTG	GCAC TGAGCT	GCGGGGACAT	6180
GAGGGCCCTG	TGAGCTGCTG	TAGTTTCAGC	ACTGATGGAG	GCAGCCTGGC	CACCGGGGGC	6240
CGGGATCGGA	GTCTCCTCTG	CTGGGACGTG	AGGACACCCA	AAACCCCTGT	TTTGATCCAC	6300
TCCTTCCCTG	CCTGTCACCG	TGACTGGTC	ACTGGCTGTG	CCTGGACCAA	AGATAACCTA	6360
CTGATATCCT	GCTCCAGTGA	TGGCTCTGTG	GGGCTCTGGG	ACCCAGAGTC	AGGACAGCGG	6420
CTTGGTCAGT	TCCTGGGTCA	TCAGAGTGCT	GTGAGCGCTG	TGGCAGCTGT	GGAGGAGCAC	6480
GTGGTGTCTG	TGAGCCGGGA	TGGGACCTTG	AAAGTGTGGG	ACCATCAAGG	CGTGGAGCTG	6540
ACCAGCATCC	CTGCTCACTC	AGGACCCATT	AGCCACTGTG	CAGCTGCCAT	GGAGCCCCGT	6600
GCAGCTGGAC	AGCCTGGTC	AGAGCTTCTG	GTGGTAACCG	TCGGGCTAGA	TGGGGCCACA	6660
CGGTTATGGC	ATCCACTCTT	GGTGTGCCAA	ACCCACACCC	TCCTGGGACA	CAGCGGCCCA	6720
GTCCGTGCTG	CTGCTGTTTC	AGAAACCTCA	GGCCTCATGC	TGACCGCCTC	TGAGGATGGT	6780
TCTGTACGGC	TCTGGCAGGT	TCCTAAGGAA	GCAGATGACA	CATGTATACC	AAGGAGTTCT	6840
GCAGCCGTCA	CTGCTGTGGC	TTGGGCACCA	GATGGTCCA	TGGCAGTATC	TGGAAATCAA	6900
GCTGGGAAAC	TAATCTTGTG	GCAGGAAGCT	AAGGCTGTGG	CCACAGCACA	GGCTCCAGGC	6960
CACATTGGTG	CTCTGATCTG	GTCCTCGGCA	CACACCTTTT	TTGTCCTCAG	TGCTGATGAG	7020
AAAATCAGCG	AGTGGCAAGT	GAAACTGCGG	AAGGGTTCGG	CACCCGGAAA	TTTGAGTCTT	7080
CACCTGAACC	GAATTCTACA	GGAGGACTTA	GGGGTGCTGA	CAAGTCTGGA	TTGGGCTCCT	7140
GATGGTCACT	TTCTCATCTT	GGCCAAAGCA	GATTGAAAGT	TACTTGCAT	GAAGCCAGGG	7200
GATGCTCCAT	CTGAAATCTG	GAGCAGCTAT	ACAGAAAATC	CTATGATATT	GTCCACCCAC	7260
AAGGAGTATG	GCATATTGT	CCTGCAGCCC	AAGGATCCTG	GAGTTCTTTC	TTTCTTGAGG	7320
CAAAAGGAAT	CAGGAGAGTT	TGAAGAGAGG	CTGAACCTTG	ATATAAACTT	AGAGAATCCT	7380
AGTAGGACCC	TAATATCGAT	AACTCAAGCC	AAACCTGAAT	CTGAGTCCTC	ATTTTGTGT	7440
GCCAGCTCTG	ATGGGATCCT	ATGGAACCTG	GCCAAATGCA	GCCCAGAAGG	AGAATGGACC	7500
ACAGGTAACA	TGTGGCAGAA	AAAAGCAAAC	ACTCCAGAAA	CCCAAACCTCC	AGGGACAGAC	7560
CCATCTACCT	GCAGGGAATC	TGATGCCAGC	ATGGATAGTG	ATGCCAGCAT	GGATAGTGAG	7620

CCAACACCAC ATCTAAAGAC ACGGCAGCGT AGAAAGATT C ACTCGGGCTC TGTCACAGCC	7680
CTCCATGTGC TACCTGAGTT GCTGGTGACA GCTTCGAAGG ACAGAGATGT TAAGCTATGG	7740
GAGAGACCCA GTATGCAGCT GCTGGGCCTG TTCCGATGCG AAGGGTCAGT GAGCTGCCTG	7800
GAACCTTGGC TGGGCGCTAA CTCCACCCTG CAGCTTGCCG TGGGAGACGT GCAGGGCAAT	7860
GTGTACTTTC TGAATTGGGA A	7881

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7886 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGGAGAACGC TCTGTGGGCA TGTGCCTGGC CATTCA GACA TCCTCTCCTT GAAGAACCGG	60
TGCCTGACCA TGCTCCCTGA CCTCCAGCCC CTGGAGAAAA TACATGGACA TAGATCTGTC	120
CACTCAGACA TCCTTTCCCTT GGAGAACCGAG TGTCTGACCA TGCTCTCTGA CCTCCAGCCC	180
ACGGAGAGAA TAGATGGGCA TATATCTGTC CACCCAGACA TCCTCTCCTT GGAGAACCGG	240
TGCCTGACCA TGCTCCCTGA CCTCCAGGCC CTGGAGAACG TATGTGGACA TATGTCTAGT	300
CATCCAGACG TCCTTTCTTT GGAAAACCAA TGTCTAGCTA CTCTCCCCAC TGTAAAGAGC	360
ACTGCATTGA CCAGCCCCCTT GCTCCAGGGT CTTCACATAT CTCATACGGC ACAAGCTGAT	420
CTGCATAGCC TGAAAACTAG CAACTGCCTG CTCCCTGAGC TTCCTACCAA GAAGACTCCA	480
TGTTTCTCTG AGGAACCTAGA CCTTCCACCT GGACCCAGGG CCCTGAAATC CATGTCTGCT	540
ACAGCTCAAG TCCAGGAAGT AGCCTTGGGT CAATGGTGTG TCTCCAAAGA AAAGGAATT	600
CAAGAAGAAG AAAGCACAGA AGTCCCAGTC CTTGTACAG TCTAAGCTTG GAAGAAGAAG	660
AAGTGGAGGC ACCGGTCTTA AAACTCACAT CTGGAGACTC TGGCTTCCAT CCTGAAACCA	720
CTGACCAGGT CCTTCAGGAG AAGAAGATGG CTCTCTTGAC CTTACTCTGC TCTGCTCTGG	780
CCTCAAATGT GAATGTGAAA GATGCATCTG ACCTTACCCG GGCATCCATC CTTGAAGTCT	840
GTAGTGCCCT GGCCTCCTTG GAACCGGAGT TCATCCTTAA GGCATCTTG TATGCTCGGC	900
AGCAACTTAA CCTCCGGGAC ATCGCCAATA CAGTTCTGGC TGTGGCTGCC CTCTTGCCAG	960
CCTGCCGCC CCATGTACGA CGGTATTACT CGGCCATTGT TCACCTGCCT TCAGACTGGA	1020
TCCAGGTAGC CGAGTTCTAC CAGAGCCTGG CAGAAGGGGA TGAGAAGAAG TTGGTGTCCC	1080
TGCCTGCCTG TCTCCGAGCT GCCATGACCG ACAAAATTGAT CGAGTTGAT GAGTACCAAGC	1140
TAGCTAAGTA CAACCCACGG AAACATCGGT CCAAGAGGCG GTCCCGCCAG CCACCCGCC	1200

CTCAAAAGAC AGAACGTCCA TTTTCAGAGA GAGGGAAATG TTTTCCAAAG AGCCTTGCG	1260
CCCTTAAAAA TGAACAGATT ACGTTGAAG CAGCTTATAA TGCAATGCCA GAGAAAACA	1320
GGCTACCACG GTTCACTCTG AAGAAGTTGG TAGAGTATCT ACATATCCAC AAGCCTGCTC	1380
AGCACGTCCA GGCCCTGCTG GGCTACAGGT ACCCAGCCAC CCTAGAGCTC TTTTCTCGGA	1440
GTCACCTCCC TGGGCCGTGG GAGTCTAGCA GAGCTGGTCA GC GGATGAAG CTCCGAAGGC	1500
CAGAGACCTG GGAGCGGGAG CTGAGTTAC GGGGAAACAA AGCTTCTGTG TGGGAGGAGC	1560
TCATAGACAA TGGGAAACTG CCCTTCATGG CCATGCTCCG GAACCTGTGT AACCTGCTGC	1620
GGACTGGGAT CAGTGCCCCGC CACCATGAAC TCGTTCTCCA GAGACTCCAG CATGAGAAAT	1680
CTGTGGTTCA CAGTCGGCAG TTTCCATTCA GATTCCCTAA TGCTCATGAC TCTATCGATA	1740
AACTTGAGGC TCAGCTCAGA AGCAAAGCAT CACCCCTTCCC TTCCAATACA ACATTGATGA	1800
AACGGATAAT GATTAGAAC TCAAAAAAAA ATAGGAGGCC TGCCAGTCGG AAGCACCTGT	1860
GCACCCCTGAC GCGCCGGCAG CTTCGGGCAG CAATGACTAT ACCTGTGATG TATGAGCAGC	1920
TCAAGCGGGA GAAACTGAGG CTGCACAAGG CCAGACAATG GAACTGTGAT GTTGAGTTGC	1980
TGGAGCGCTA TCGCCAGGCC CTGGAAACAG CTGTGAACCT CTCAGTAAAG CACAACCTAT	2040
CCCCGATGCC TGGCCGAACC CTCTGGTCT ATCTCACAGA TGCAAATGCC GACAGGCTCT	2100
GTCCCAAGAG TCACTCACAA GGGCCTCCCC TGAACTATGT GCTGCTGCTG ATCGGAATGA	2160
TGGTGGCTCG AGCCGAGCAA GTGACTGTT GCTTGTGTGG GGGAGGATTG GTGAAGACAC	2220
CGGTACTTAC AGCCGATGAA GGCATCCTGA AGACTGCCAT CAAACTTCAG GCTCAAGTCC	2280
AGGAGTTAGA AGGCAATGAT GAGTGGCCCC TGGACACTTT TGGGAAGTAT CTGCTGTCTC	2340
TGGCTGTCCA AAGGACCCCC ATTGACAGGG TCATCCTGTT TGGTCAAAGG ATGGATACCG	2400
AGCTCCTGAA AGTAGCCAAA CAGATTATCT GGCAGCATGT GAATTCCAAG TGCCTCTTG	2460
TTGGTGTCCCT CCTACAGAAA ACACAGTACA TATCACCAAA TTTGAATCCC AACGATGTGA	2520
CGCTCTCAGG CTGCACTGAC GGGATCCTGA AATTCAATTGC CGAACATGGA GCCTCTCGTC	2580
TCCTGGAACA TGTGGACAA CTAGATAAAC TATTCAAGAT CCCCCCACCC CCAGGAAAGA	2640
CACAGGCACC GTCTCTCCGG CCGCTGGAGG AGAACATCCC TGGTCCCTTG GGTCTATT	2700
CCCAGCATGG ATGGCGCAAT ATCCGGCTTT TCATTTCATC CACTTCCGT GACATGCATG	2760
GGGAGCGAGA TTTGCTGATG AGATCTGTT TGCCCGCACT GCAGGCCAGA GTGTTCCCCC	2820
ACCGCATCAG TCTTCACGCC ATTGACCTGC GCTGGGGTAT CACAGAGGAA GAGACCCGCA	2880
GGAACAGACA ACTGGAAGTG TGCCTTGGGG AGGTGGAGAA CTCACAGCTG TTCGTGGGA	2940
TTCTGGGCTC CCGCTATGGC TACATTCCCC CCAGCTATGA TCTTCCTGAT CATCCCCACT	3000
TTCACTGGAC CCATGAGTAC CCTTCAGGGC GATCCGTGAC AGAGATGGAG GTGATGCAAT	3060

TCCCTGAACCG	TGGCCAACGC	TCGCAGCCTT	CGGCCAAGC	TCTCATCTAC	TTCCGAGATC	3120
CTGATTCCT	TAGCTCTGTG	CCAGATGCCT	GGAAACCTGA	CTTTATATCT	GAGTCAGAACG	3180
AAGCTGCACA	TCGGGTCTCA	GAGCTGAAGA	GATATCTACA	CGAACAGAAA	GAGGTTACCT	3240
GTCGCAGCTA	CTCCTGTGAA	TGGGGAGGTG	TAGCGGCTGG	CCGGCCCTAT	ACTGGGGGCC	3300
TGGAGGAGTT	TGGACAGTTG	GTTCTCCAGG	ATGTGTGGAG	CATGATCCAG	AAGCAGCACC	3360
TGCAGCCTGG	GGCCCAGTTG	GAGCAGCCAA	CATCCATCTC	AGAAGACGAT	TTGATCCAGA	3420
CCAGCTTCA	GCAGCTGAAG	ACCCCAACGA	GTCCGGCACG	GCCACGCCTT	CTTCAGGATA	3480
CAGTGCAGCA	GCTGTTGCTG	CCCCATGGGA	GGCTGAGCCT	AGTGACTGGG	CAGGCAGGAC	3540
AGGGAAAGAC	TGCCTTCTG	GCATCCCTTG	TGTCTGCCCT	GAAGGTCCCT	GACCAGCCCA	3600
ATGAGCCCCC	GTTCGTTTTC	TTCCACTTTG	CAGCAGCCCG	CCCTGACCAG	TGTCTGCTC	3660
TCAACCTCCT	CAGACGCCTC	TGTACCCATC	TGCGTAAAAA	ACTGGGAGAG	CTGAGTGCCC	3720
TCCCCAGCAC	TTACAGAGGC	CTGGTGTGGG	AACTGCAGCA	GAAGTTGCTC	CTCAAATTG	3780
CTCAGTCGCT	GCAGCCTGCT	CAGACTTTGG	TCCTTATCAT	CGATGGGGCA	GATAAGTTGG	3840
TGGATCGTAA	TGGGCAGCTG	ATTCAGACT	GGATCCCCAA	GTCTCTTCCG	CGGCGAGTAC	3900
ACCTGGTGCT	GAGTGTGTCC	AGTGACTCAG	GCCTGGGTGA	GACCCTTCAG	CAAAGTCAGG	3960
GTGCTTATGT	GGTGGCCTTG	GGCTCTTGG	TCCCACATTC	AAGGGCTCAG	CTTGTGAGAG	4020
AAGAGCTAGC	ACTGTATGGG	AAACGACTGG	AGGAGTCACC	TTTAACAAAC	CAGATGCGGC	4080
TGCTGCTGGC	AAAGCAGGGT	TCAAGCCTGC	CATTGTACCT	GCACCTTGTC	ACTGACTACC	4140
TGAGGCTCTT	CACACTGTAT	GAACAGGTGT	CTGAGAGACT	TCGAACCCCTG	CCCGGCCACTC	4200
TCCCCACTGCT	CTTGCAGCAC	ATCCTGAGCA	CCTTGGAGCA	AGAACATGGC	CATGATGTCC	4260
TTCCTCAGGC	TTTGACTGCC	CTTGAGGTCA	CACGAAGTGG	TCTGACTGTG	GACCAGCTAC	4320
ATGCAATCCT	GAGCACATGG	CTGATCTTC	CCAAGGAGAC	TAAGAGCTGG	GAAGAAGTGC	4380
TGGCTGCCAG	TCACAGTGGA	AACCCTTCC	CCTTGTGTCC	ATTTGCCTAC	CTTGTCCAGA	4440
GTCTACGCAG	TTTACTAGGG	GAGGGCCAG	TGGAGCGCCC	TGGTGCCCGT	CTCTGCCTCT	4500
CTGATGGGCC	CCTGAGGACA	ACAATTAAAC	GTCGCTATGG	GAAAAGGCTG	GGGCTAGAGA	4560
AGACTGCGCA	TGTCCCTCATT	GCAGCTCACC	TCTGGAAGAC	GTGTGATCCT	GATGCCTCGG	4620
GCACCTTCCG	AAAGTGCCT	CCTGAGGCTC	TGAAAGATTT	ACCTTACAC	CTGCTCCAGA	4680
GCGGGAAACCA	TGGTCTCCTT	GCCGAGTTTC	TTACCAATCT	CCATGTGGTT	GCTGCATATC	4740
TGGAAGTGGG	TCTAGTCCCC	GACCTTGG	AGGCTCATGT	GCTCTATGCT	TCTTCAAAGC	4800
CTGAAGCCAA	CCAGAAGCTC	CCAGCGGCAG	ATGTTGCTGT	TTTCCATACC	TTCCTGAGAC	4860
AACAGGCTTC	ACTCCTTACC	CAGTATCCTT	TGCTCCTGCT	CCAGCAGGCA	GCTAGCCAGC	4920
CTGAAGAGTC	ACCTGTTGC	TGCCAGGGCCC	CCCTGCTCAC	CCAGCGATGG	CACGACCAGT	4980

TCACACTGAA ATGGATTAAT AAACCCCAGA CCCTGAAGGG TCAGCAAAGC TTGTCTCTGA	5040
CAATGTCCCTC ATCCCCAACT GCTGTGGCCT TCTCCCCGAA TGGGCAAAGA GCAGCTGTGG	5100
GGACCGCCAG TGGGACAATT TACCTGTTGA ACTTGAAAAC CTGGCAGGAG GAGAAGGCTG	5160
TGGTGAGTGG CTGTGACGGG ATTTCCTCTT TTGCATTCCCT TTCGGACACT GCCCTTTCC	5220
TTACTACCTT CGACGGGCAC CTAGAGCTTT GGGACCTGCA ACATGGTTGT TGGGTGTTTC	5280
AGACCAAGGC CCACCAAGTAC CAAATCACTG GCTGCTGCCT GAGCCCAGAC CGCCGCCCTGC	5340
TGGCCACTGT GTGTTGGGA GGATACCTAA AGCTGTGGGA CACAGTCCGA GGACAGCTGG	5400
CTTTTCAGTA CACCCATCCA AAGTCTCTCA ACTGCGTTGC CTTCCACCCA GAGGGGCAGG	5460
TGGTAGCCAC AGGCAGCTGG GCTGGCAGCA TTACCTTCTT CCAGGCAGAT GGACTCAAAG	5520
TCACCAAGGA ACTAGGGGCC CCCGGACCCCT CTGCTGTAG TTTGGCATTC AACAAACCTG	5580
GGAAGATTGT GGCTGTGGC CGGATAGATG GGACAGTGGA GCTGTGGCC TGGCAAGAGG	5640
GTGCCCGGCT GGCGGCCTTC CCTGCACAGT GTGGCTGTGT CTCTGCTGTT CTTTCTTGC	5700
ATGCTGGAGA CCGGTTCTG ACTGCTGGAG AAGATGGCAA GGCTCAGTTA TGGTCAGGAT	5760
TTCTTGGCCG GCCCAGGGGT TGCCTGGGCT CTCTTCCTCT TTCTCCTGCA CTCTCGGTGG	5820
CTCTCAACCC AGACGGTGAC CAGGTGGCTG TTGGGTACCG AGAAGATGGC ATTAACATCT	5880
ACAAGATTTC TTCAGGTTCC CAGGGGCCTC AGCATCAAGA GCTAAATGTG GCGGTGTCTG	5940
CACTGGTGTG GCTGAGCCCT AGTGTGTTGG TGAGTGGTGC AGAAGATGGA TCCCTGCATG	6000
GTTGGATGTT CAAGGGAGAC TCCCTTCATT CCCTGTGGCT GTTGTGAGA TACCAGAAC	6060
CTGTGCTGGG ACTGGCTGCC TCCCAGGAAC TCATGGCTGC TGCCTCAGAG GACTTCAC	6120
TGAGACTGTG GCCCAGACAG CTGCTGACAC AGCCACATGT GCATGCGGTA GAGTTGCCCT	6180
GTTGTGCTGA ACTCCGGGGA CACGAGGGC CAGTGTGCTG CTGTAGCTTC AGCCCTGATG	6240
GAGGCATCTT GGCCACAGCT GGCAGGGATC GGAATCTCCT TTGCTGGAC ATGAAGATA	6300
CCCAAGCCCC TCTCCTGATT CACACTTTCT CGTCCTGTCA TCGTGA	6360
CTGGACTGGAC CAAAGACAAC ATCCTGGTCT CCTGCTCGAG TGATGGCTCT GTGGGACTCT	6420
GGAACCCAGA GGCAGGGCAG CAACTTGGCC AGTTCTCAGG CCACCAGAGT GCCGTGAGCG	6480
CCGTGGTTGC TGTGGAGGAA CACATTGTAT CTGTGAGCCG AGATGGGACC TTGAAAGTGT	6540
GGGACCATCA GGGTGTGGAG CTGACCAGCA TCCCTGCCA TTCCGGACCC ATCAGCCAGT	6600
GTGCAGCTGC TCTGGAGCCC CGCCCAGGGG GACAGCCTGG ATCAGAGCTT CTGGTGGTGA	6660
CTGTTGGACT AGATGGGCC ACAAAAGTTGT GGCATCCCCT GTTGGTGTGC CAAATACGTA	6720
CTCTCCAGGG ACACAGTGGC CCAGTCACAG CAGCTGCTGC TTCAGAGGCC TCAGGCCTCC	6780
TGCTGACCTC AGATGATAGC TCTGTACAGC TCTGGCAGAT ACCAAAGGAA CCAGATGATT	6840

CATAACAAACC TAGGAGTTCT GTGGCCATCA CTGCTGTGGC ATGGGCACCG GATGGTTCTA	6900
TGGTGGTGTGTC CGGAAATGAA GCCGGGAAAC TGACACTGTG GCAGCAAGCC AAGGCTGTGG	6960
CTACCGCACA GGCTCCAGGC CGCGTCAGTC ACCTGATCTG GTACTCGGCA AATTCAATTCT	7020
TCGTTCTCAG TGCTAATGAA AACGTCAGCG AGTGGCAAGT GGGACTGAGG AAAGGTTCAA	7080
CGTCCACCAG TTCCAGTCTT CATCTGAAGA GAGTTCTGCA GGAGGACTGG GGAGTCTTGA	7140
CAGGTCTGGG TCTGGCCCCT GATGGCCAGT CTCTCATCTT GATGAAAGAG GATGTGGAAT	7200
TACTAGAGAT GAAGCCTGGG TCTATTCCAT CTTCTATCTG CAGGAGGTAT GGAGTACATT	7260
CTTCAATACT GTGCACCAGC AAGGAGTACG GCTTGTCTA CCTGCAGCAG GGGGACTCCG	7320
GATTACTTTC TATATTGGAG CAAAAGGAGT CAGGGGAGTT TGAAGAGATC CTGGACTTCA	7380
ATCTGAACCT AAATAATCCT AATGGGTCCC CAGTATCAAT CACTCAGGCC AAACCTGAGT	7440
CTGAATCATC CCTTTTGTGC GCCACCTCTG ATGGGATGCT GTGGAACCTTA TCTGAATGTA	7500
CCTCAGAGGG AGAATGGATC GTAGATAACA TTTGGCAGAA AAAAGCAAAA AAACCTAAAA	7560
CTCAGACTCT GGAGACAGAG TTGTCCCCGC ACTCAGAGTT GGATTTTCC ATTGATTGCT	7620
GGATTGATCC CACAAATTAA AAGGCACAGC AGTGTAAAAA GATCCACTTG GGCTCTGTCA	7680
CAGCCCTCCA TGTGCTTCCG GGATTGCTGG TGACAGCTTC GAAGGACAGA GATGTTAAC	7740
TGTGGGAGAG ACCCAGTATG CAGCTGCTGG GCTTGTCTCG ATGTGAAGGG CCAGTGAGCT	7800
GTCTGGAACC TTGGATGGAG CCCAGCTCTC CCCTGCAGCT TGCTGTGGGAA GACACACAAG	7860
GAAACTTGTAA TTTTCTATCT TGGGAA	7886

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2627 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Glu Lys Leu His Gly His Val Ser Ala His Pro Asp Ile Leu Ser			
1	5	10	15
Leu Glu Asn Arg Cys Leu Ala Met Leu Pro Asp Leu Gln Pro Leu Glu			
20	25	30	
Lys Leu His Gln His Val Ser Thr His Ser Asp Ile Leu Ser Leu Lys			
35	40	45	
Asn Gln Cys Leu Ala Thr Leu Pro Asp Leu Lys Thr Met Glu Lys Pro			
50	55	60	
His Gly Tyr Val Ser Ala His Pro Asp Ile Leu Ser Leu Glu Asn Gln			
65	70	75	80

Cys Leu Ala Thr Leu Ser Asp Leu Lys Thr Met Glu Lys Pro His Gly
85 90 95

His Val Ser Ala His Pro Asp Ile Leu Ser Leu Glu Asn Arg Cys Leu
100 105 110

Ala Thr Leu Pro Ser Leu Lys Ser Thr Val Ser Ala Ser Pro Leu Phe
115 120 125

Gln Ser Leu Gln Ile Ser His Met Thr Gln Ala Asp Leu Tyr Arg Val
130 135 140

Asn Asn Ser Asn Cys Leu Leu Ser Glu Pro Pro Ser Trp Arg Ala Gln
145 150 155 160

His Phe Ser Lys Gly Leu Asp Leu Ser Thr Cys Pro Ile Ala Leu Lys
165 170 175

Ser Ile Ser Ala Thr Glu Thr Ala Gln Glu Ala Thr Leu Gly Arg Trp
180 185 190

Phe Asp Ser Glu Glu Lys Lys Gly Ala Glu Thr Gln Met Pro Ser Tyr
195 200 205

Ser Leu Ser Leu Gly Glu Glu Glu Val Glu Asp Leu Ala Val Lys
210 215 220

Leu Thr Ser Gly Asp Ser Glu Ser His Pro Glu Pro Thr Asp His Val
225 230 235 240

Leu Gln Glu Lys Lys Met Ala Leu Leu Ser Leu Leu Cys Ser Thr Leu
245 250 255

Val Ser Glu Val Asn Met Asn Asn Thr Ser Asp Pro Thr Leu Ala Ala
260 265 270

Ile Phe Glu Ile Cys Arg Glu Leu Ala Leu Leu Glu Pro Glu Phe Ile
275 280 285

Leu Lys Ala Ser Leu Tyr Ala Arg Gln Gln Leu Asn Val Arg Asn Val
290 295 300

Ala Asn Asn Ile Leu Ala Ile Ala Ala Phe Leu Pro Ala Cys Arg Pro
305 310 315 320

His Leu Arg Arg Tyr Phe Cys Ala Ile Val Gln Leu Pro Ser Asp Trp
325 330 335

Ile Gln Val Ala Glu Leu Tyr Gln Ser Leu Ala Glu Gly Asp Lys Asn
340 345 350

Lys Leu Val Pro Leu Pro Ala Cys Leu Arg Thr Ala Met Thr Asp Lys
355 360 365

Phe Ala Gln Phe Asp Glu Tyr Gln Leu Ala Lys Tyr Asn Pro Arg Lys
370 375 380

His Arg Ala Lys Arg His Pro Arg Arg Pro Pro Arg Ser Pro Gly Met
385 390 395 400

Glu Pro Pro Phe Ser His Arg Cys Phe Pro Arg Tyr Ile Gly Phe Leu
405 410 415

Arg Glu Glu Gln Arg Lys Phe Glu Lys Ala Gly Asp Thr Val Ser Glu
420 425 430

Lys Lys Asn Pro Pro Arg Phe Thr Leu Lys Lys Leu Val Gln Arg Leu
435 440 445

His Ile His Lys Pro Ala Gln His Val Gln Ala Leu Leu Gly Tyr Arg
450 455 460

Tyr Pro Ser Asn Leu Gln Leu Phe Ser Arg Ser Arg Leu Pro Gly Pro
465 470 475 480

Trp Asp Ser Ser Arg Ala Gly Lys Arg Met Lys Leu Ser Arg Pro Glu
485 490 495

Thr Trp Glu Arg Glu Leu Ser Leu Arg Gly Asn Lys Ala Ser Val Trp
500 505 510

Glu Glu Leu Ile Glu Asn Gly Lys Leu Pro Phe Met Ala Met Leu Arg
515 520 525

Asn Leu Cys Asn Leu Leu Arg Val Gly Ile Ser Ser Arg His His Glu
530 535 540

Leu Ile Leu Gln Arg Leu Gln His Gly Lys Ser Val Ile His Ser Arg
545 550 555 560

Gln Phe Pro Phe Arg Phe Leu Asn Ala His Asp Ala Ile Asp Ala Leu
565 570 575

Glu Ala Gln Leu Arg Asn Gln Ala Leu Pro Phe Pro Ser Asn Ile Thr
580 585 590

Leu Met Arg Arg Ile Leu Thr Arg Asn Glu Lys Asn Arg Pro Arg Arg
595 600 605

Arg Phe Leu Cys His Leu Ser Arg Gln Gln Leu Arg Met Ala Met Arg
610 615 620

Ile Pro Val Leu Tyr Glu Gln Leu Lys Arg Glu Lys Leu Arg Val His
625 630 635 640

Lys Ala Arg Gln Trp Lys Tyr Asp Gly Glu Met Leu Asn Arg Tyr Arg
645 650 655

Gln Ala Leu Glu Thr Ala Val Asn Leu Ser Val Lys His Ser Leu Pro
660 665 670

Leu Leu Pro Gly Arg Thr Val Leu Val Tyr Leu Thr Asp Ala Asn Ala
675 680 685

Asp Arg Leu Cys Pro Lys Ser Asn Pro Gln Gly Pro Pro Leu Asn Tyr
690 695 700

Ala Leu Leu Leu Ile Gly Met Met Ile Thr Arg Ala Glu Gln Val Asp
705 710 715 720

Val Val Leu Cys Gly Gly Asp Thr Leu Lys Thr Ala Val Leu Lys Ala
725 730 735

Glu Glu Gly Ile Leu Lys Thr Ala Ile Lys Leu Gln Ala Gln Val Gln
740 745 750

Glu Phe Asp Glu Asn Asp Gly Trp Ser Leu Asn Thr Phe Gly Lys Tyr
 755 760 765
 Leu Leu Ser Leu Ala Gly Gln Arg Val Pro Val Asp Arg Val Ile Leu
 770 775 780
 Leu Gly Gln Ser Met Asp Asp Gly Met Ile Asn Val Ala Lys Gln Leu
 785 790 795 800
 Tyr Trp Gln Arg Val Asn Ser Lys Cys Leu Phe Val Gly Ile Leu Leu
 805 810 815
 Arg Arg Val Gln Tyr Leu Ser Thr Asp Leu Asn Pro Asn Asp Val Thr
 820 825 830
 Leu Ser Gly Cys Thr Asp Ala Ile Leu Lys Phe Ile Ala Glu His Gly
 835 840 845
 Ala Ser His Leu Leu Glu His Val Gly Gln Met Asp Lys Ile Phe Lys
 850 855 860
 Ile Pro Pro Pro Pro Gly Lys Thr Gly Val Gln Ser Leu Arg Pro Leu
 865 870 875 880
 Glu Glu Asp Thr Pro Ser Pro Leu Ala Pro Val Ser Gln Gln Gly Trp
 885 890 895
 Arg Ser Ile Arg Leu Phe Ile Ser Ser Thr Phe Arg Asp Met His Gly
 900 905 910
 Glu Arg Asp Leu Leu Leu Arg Ser Val Leu Pro Ala Leu Gln Ala Arg
 915 920 925
 Ala Ala Pro His Arg Ile Ser Leu His Gly Ile Asp Leu Arg Trp Gly
 930 935 940
 Val Thr Glu Glu Glu Thr Arg Arg Asn Arg Gln Leu Glu Val Cys Leu
 945 950 955 960
 Gly Glu Val Glu Asn Ala Gln Leu Phe Val Gly Ile Leu Gly Ser Arg
 965 970 975
 Tyr Gly Tyr Ile Pro Pro Ser Tyr Asn Leu Pro Asp His Pro His Phe
 980 985 990
 His Trp Ala Gln Gln Tyr Pro Ser Gly Arg Ser Val Thr Glu Met Glu
 995 1000 1005
 Val Met Gln Phe Leu Asn Arg Asn Gln Arg Leu Gln Pro Ser Ala Gln
 1010 1015 1020
 Ala Leu Ile Tyr Phe Arg Asp Ser Ser Phe Leu Ser Ser Val Pro Asp
 1025 1030 1035 1040
 Ala Trp Lys Ser Asp Phe Val Ser Glu Ser Glu Glu Ala Ala Xaa Arg
 1045 1050 1055
 Ile Ser Glu Leu Lys Ser Tyr Leu Ser Arg Gln Lys Gly Ile Thr Cys
 1060 1065 1070
 Arg Arg Tyr Pro Cys Glu Trp Gly Gly Val Ala Ala Gly Arg Pro Tyr
 1075 1080 1085

Val Gly Gly Leu Glu Glu Phe Gly Gln Leu Val Leu Gln Asp Val Trp
1090 1095 1100

Asn Met Ile Gln Lys Leu Tyr Leu Gln Pro Gly Ala Leu Leu Glu Gln
1105 1110 1115 1120

Pro Val Ser Ile Pro Asp Asp Asp Leu Val Gln Ala Thr Phe Gln Gln
1125 1130 1135

Leu Gln Lys Pro Pro Ser Pro Ala Arg Pro Arg Leu Leu Gln Asp Thr
1140 1145 1150

Val Gln Xaa Leu Met Leu Pro His Gly Arg Leu Ser Leu Val Thr Gly
1155 1160 1165

Gln Ser Gly Gln Gly Lys Thr Ala Phe Leu Ala Ser Leu Val Ser Ala
1170 1175 1180

Leu Gln Ala Pro Asp Gly Ala Lys Val Ala Xaa Leu Val Phe Phe His
1185 1190 1195 1200

Phe Ser Gly Ala Arg Pro Asp Gln Gly Leu Ala Leu Thr Leu Leu Arg
1205 1210 1215

Arg Leu Cys Thr Tyr Leu Arg Gly Gln Leu Lys Glu Pro Gly Ala Leu
1220 1225 1230

Pro Ser Thr Tyr Arg Ser Leu Val Trp Glu Leu Gln Gln Arg Leu Leu
1235 1240 1245

Pro Lys Ser Ala Glu Ser Leu His Pro Gly Gln Thr Gln Val Leu Ile
1250 1255 1260

Ile Asp Gly Ala Asp Arg Leu Val Asp Gln Asn Gly Gln Leu Ile Ser
1265 1270 1275 1280

Asp Trp Ile Pro Lys Lys Leu Pro Arg Cys Val His Leu Val Leu Ser
1285 1290 1295

Val Ser Ser Asp Ala Gly Leu Gly Glu Thr Leu Glu Gln Ser Gln Gly
1300 1305 1310

Ala His Val Leu Ala Leu Gly Pro Leu Glu Ala Ser Ala Arg Ala Arg
1315 1320 1325

Leu Val Arg Glu Glu Leu Ala Leu Tyr Gly Lys Arg Leu Glu Glu Ser
1330 1335 1340

Pro Phe Asn Asn Gln Met Arg Leu Leu Val Lys Arg Glu Ser Gly
1345 1350 1355 1360

Arg Pro Leu Tyr Leu Arg Leu Val Thr Asp His Leu Arg Leu Phe Thr
1365 1370 1375

Leu Tyr Glu Gln Val Ser Glu Arg Leu Arg Thr Leu Pro Ala Thr Val
1380 1385 1390

Pro Leu Leu Leu Gln His Ile Leu Ser Thr Leu Glu Lys Glu His Gly
1395 1400 1405

Pro Asp Val Leu Pro Gln Ala Leu Thr Ala Leu Glu Val Thr Arg Ser
1410 1415 1420

Gly Leu Thr Val Asp Gln Leu His Gly Val Leu Ser Val Trp Arg Thr
1425 1430 1435 1440

Leu Pro Lys Gly Thr Lys Ser Trp Glu Glu Ala Val Ala Ala Gly Asn
1445 1450 1455

Ser Gly Asp Pro Tyr Pro Met Gly Pro Phe Ala Cys Leu Val Gln Ser
1460 1465 1470

Leu Arg Ser Leu Leu Gly Glu Gly Pro Leu Glu Arg Pro Gly Ala Arg
1475 1480 1485

Leu Cys Leu Pro Asp Gly Pro Leu Arg Thr Ala Ala Lys Arg Cys Tyr
1490 1495 1500

Gly Lys Arg Pro Gly Leu Glu Asp Thr Ala His Ile Leu Ile Ala Ala
1505 1510 1515 1520

Gln Leu Trp Lys Thr Cys Asp Ala Asp Ala Ser Gly Thr Phe Arg Ser
1525 1530 1535

Cys Pro Pro Glu Ala Leu Gly Asp Leu Pro Tyr His Leu Leu Gln Ser
1540 1545 1550

Gly Asn Arg Gly Leu Leu Ser Lys Phe Leu Thr Asn Leu His Val Val
1555 1560 1565

Ala Ala His Leu Glu Leu Gly Leu Val Ser Arg Leu Leu Glu Ala His
1570 1575 1580

Ala Leu Tyr Ala Ser Ser Val Pro Lys Glu Glu Gln Lys Leu Pro Glu
1585 1590 1595 1600

Ala Asp Val Ala Val Phe Arg Thr Phe Leu Arg Gln Gln Ala Ser Ile
1605 1610 1615

Leu Ser Gln Tyr Pro Arg Leu Leu Pro Gln Gln Ala Ala Asn Gln Pro
1620 1625 1630

Leu Asp Ser Pro Leu Cys His Gln Ala Ser Leu Leu Ser Arg Arg Trp
1635 1640 1645

His Leu Gln His Thr Leu Arg Trp Leu Asn Lys Pro Arg Thr Met Lys
1650 1655 1660

Asn Gln Gln Ser Ser Ser Leu Ser Leu Ala Val Ser Ser Ser Pro Thr
1665 1670 1675 1680

Ala Val Ala Phe Ser Thr Asn Gly Gln Arg Ala Ala Val Gly Thr Ala
1685 1690 1695

Asn Gly Thr Val Tyr Leu Leu Asp Leu Arg Thr Trp Gln Glu Glu Lys
1700 1705 1710

Ser Val Val Ser Gly Cys Asp Gly Ile Ser Ala Cys Leu Phe Leu Ser
1715 1720 1725

Asp Asp Thr Leu Phe Leu Thr Ala Phe Asp Gly Leu Leu Glu Leu Trp
1730 1735 1740

Asp Leu Gln His Gly Cys Arg Val Leu Gln Thr Lys Ala His Gln Tyr
1745 1750 1755 1760

Gln Ile Thr Gly Cys Cys Leu Ser Pro Asp Cys Arg Leu Leu Ala Thr
1765 1770 1775

Val Cys Leu Gly Gly Cys Leu Lys Leu Trp Asp Thr Val Arg Gly Gln
1780 1785 1790

Leu Ala Phe Gln His Thr Tyr Pro Lys Ser Leu Asn Cys Val Ala Phe
1795 1800 1805

His Pro Glu Gly Gln Val Ile Ala Thr Gly Ser Trp Ala Gly Ser Ile
1810 1815 1820

Ser Phe Phe Gln Val Asp Gly Leu Lys Val Thr Lys Asp Leu Gly Ala
1825 1830 1835 1840

Pro Gly Ala Ser Ile Arg Thr Leu Ala Phe Asn Val Pro Gly Gly Val
1845 1850 1855

Val Ala Val Gly Arg Leu Asp Ser Met Val Glu Leu Trp Ala Trp Arg
1860 1865 1870

Glu Gly Ala Arg Leu Ala Ala Phe Pro Ala His His Gly Phe Val Ala
1875 1880 1885

Ala Ala Leu Phe Leu His Ala Gly Cys Gln Leu Leu Thr Ala Gly Glu
1890 1895 1900

Asp Gly Lys Val Gln Val Trp Ser Gly Ser Leu Gly Arg Pro Arg Gly
1905 1910 1915 1920

His Leu Gly Ser Leu Ser Leu Ser Pro Ala Leu Ser Val Ala Leu Ser
1925 1930 1935

Pro Asp Gly Asp Arg Val Ala Val Gly Tyr Arg Ala Asp Gly Ile Arg
1940 1945 1950

Ile Tyr Lys Ile Ser Ser Gly Ser Gln Gly Ala Gln Gly Gln Ala Leu
1955 1960 1965

Asp Val Ala Val Ser Ala Leu Ala Trp Leu Ser Pro Lys Val Leu Val
1970 1975 1980

Ser Gly Ala Glu Asp Gly Ser Leu Gln Gly Trp Ala Leu Lys Glu Cys
1985 1990 1995 2000

Ser Leu Gln Ser Leu Trp Leu Leu Ser Arg Phe Gln Lys Pro Val Leu
2005 2010 2015

Gly Leu Ala Thr Ser Gln Glu Leu Leu Ala Ser Ala Ser Glu Asp Phe
2020 2025 2030

Thr Val Gln Leu Trp Pro Arg Gln Leu Leu Thr Arg Pro His Lys Ala
2035 2040 2045

Glu Asp Phe Pro Cys Gly Thr Glu Leu Arg Gly His Glu Gly Pro Val
2050 2055 2060

Ser Cys Cys Ser Phe Ser Thr Asp Gly Gly Ser Leu Ala Thr Gly Gly
2065 2070 2075 2080

Arg Asp Arg Ser Leu Leu Cys Trp Asp Val Arg Thr Pro Lys Thr Pro
2085 2090 2095

Val Leu Ile His Ser Phe Pro Ala Cys His Arg Asp Trp Val Thr Gly
2100 2105 2110

Cys Ala Trp Thr Lys Asp Asn Leu Leu Ile Ser Cys Ser Ser Asp Gly
2115 2120 2125

Ser Val Gly Leu Trp Asp Pro Glu Ser Gly Gln Arg Leu Gly Gln Phe
2130 2135 2140

Leu Gly His Gln Ser Ala Val Ser Ala Val Ala Val Glu Glu His
2145 2150 2155 2160

Val Val Ser Val Ser Arg Asp Gly Thr Leu Lys Val Trp Asp His Gln
2165 2170 2175

Gly Val Glu Leu Thr Ser Ile Pro Ala His Ser Gly Pro Ile Ser His
2180 2185 2190

Cys Ala Ala Ala Met Glu Pro Arg Ala Ala Gly Gln Pro Gly Ser Glu
2195 2200 2205

Leu Leu Val Val Thr Val Gly Leu Asp Gly Ala Thr Arg Leu Trp His
2210 2215 2220

Pro Leu Leu Val Cys Gln Thr His Thr Leu Leu Gly His Ser Gly Pro
2225 2230 2235 2240

Val Arg Ala Ala Ala Val Ser Glu Thr Ser Gly Leu Met Leu Thr Ala
2245 2250 2255

Ser Glu Asp Gly Ser Val Arg Leu Trp Gln Val Pro Lys Glu Ala Asp
2260 2265 2270

Asp Thr Cys Ile Pro Arg Ser Ser Ala Ala Val Thr Ala Val Ala Trp
2275 2280 2285

Ala Pro Asp Gly Ser Met Ala Val Ser Gly Asn Gln Ala Gly Glu Leu
2290 2295 2300

Ile Leu Trp Gln Glu Ala Lys Ala Val Ala Thr Ala Gln Ala Pro Gly
2305 2310 2315 2320

His Ile Gly Ala Leu Ile Trp Ser Ser Ala His Thr Phe Phe Val Leu
2325 2330 2335

Ser Ala Asp Glu Lys Ile Ser Glu Trp Gln Val Lys Leu Arg Lys Gly
2340 2345 2350

Ser Ala Pro Gly Asn Leu Ser Leu His Leu Asn Arg Ile Leu Gln Glu
2355 2360 2365

Asp Leu Gly Val Leu Thr Ser Leu Asp Trp Ala Pro Asp Gly His Phe
2370 2375 2380

Leu Ile Leu Ala Lys Ala Asp Leu Lys Leu Leu Cys Met Lys Pro Gly
2385 2390 2395 2400

Asp Ala Pro Ser Glu Ile Trp Ser Ser Tyr Thr Glu Asn Pro Met Ile
2405 2410 2415

Leu Ser Thr His Lys Glu Tyr Gly Ile Phe Val Leu Gln Pro Lys Asp
2420 2425 2430

Pro Gly Val Leu Ser Phe Leu Arg Gln Lys Glu Ser Gly Glu Phe Glu
 2435 2440 2445
 Glu Arg Leu Asn Phe Asp Ile Asn Leu Glu Asn Pro Ser Arg Thr Leu
 2450 2455 2460
 Ile Ser Ile Thr Gln Ala Lys Pro Glu Ser Glu Ser Ser Phe Leu Cys
 2465 2470 2475 2480
 Ala Ser Ser Asp Gly Ile Leu Trp Asn Leu Ala Lys Cys Ser Pro Glu
 2485 2490 2495
 Gly Glu Trp Thr Thr Gly Asn Met Trp Gln Lys Lys Ala Asn Thr Pro
 2500 2505 2510
 Glu Thr Gln Thr Pro Gly Thr Asp Pro Ser Thr Cys Arg Glu Ser Asp
 2515 2520 2525
 Ala Ser Met Asp Ser Asp Ala Ser Met Asp Ser Glu Pro Thr Pro His
 2530 2535 2540
 Leu Lys Thr Arg Gln Arg Arg Lys Ile His Ser Gly Ser Val Thr Ala
 2545 2550 2555 2560
 Leu His Val Leu Pro Glu Leu Leu Val Thr Ala Ser Lys Asp Arg Asp
 2565 2570 2575
 Val Lys Leu Trp Glu Arg Pro Ser Met Gln Leu Leu Gly Leu Phe Arg
 2580 2585 2590
 Cys Glu Gly Ser Val Ser Cys Leu Glu Pro Trp Leu Gly Ala Asn Ser
 2595 2600 2605
 Thr Leu Gln Leu Ala Val Gly Asp Val Gln Gly Asn Val Tyr Phe Leu
 2610 2615 2620
 Asn Trp Glu
 2625

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2629 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu Lys Leu Cys Gly His Val Pro Gly His Ser Asp Ile Leu Ser
 1 5 10 15
 Leu Lys Asn Arg Cys Leu Thr Met Leu Pro Asp Leu Gln Pro Leu Glu
 20 25 30
 Lys Ile His Gly His Arg Ser Val His Ser Asp Ile Leu Ser Leu Glu
 35 40 45

Asn Gln Cys Leu Thr Met Leu Ser Asp Leu Gln Pro Thr Glu Arg Ile
 50 55 60
 Asp Gly His Ile Ser Val His Pro Asp Ile Leu Ser Leu Glu Asn Arg
 65 70 75 80
 Cys Leu Thr Met Leu Pro Asp Leu Gln Pro Leu Glu Lys Leu Cys Gly
 85 90 95
 His Met Ser Ser His Pro Asp Val Leu Ser Leu Glu Asn Gln Cys Leu
 100 105 110
 Ala Thr Leu Pro Thr Val Lys Ser Thr Ala Leu Thr Ser Pro Leu Leu
 115 120 125
 Gln Gly Leu His Ile Ser His Thr Ala Gln Ala Asp Leu His Ser Leu
 130 135 140
 Lys Thr Ser Asn Cys Leu Leu Pro Glu Leu Pro Thr Lys Lys Thr Pro
 145 150 155 160
 Cys Phe Ser Glu Glu Leu Asp Leu Pro Pro Gly Pro Arg Ala Leu Lys
 165 170 175
 Ser Met Ser Ala Thr Ala Gln Val Gln Glu Val Ala Leu Gly Gln Trp
 180 185 190
 Cys Val Ser Lys Glu Lys Glu Phe Gln Glu Glu Glu Ser Thr Glu Val
 195 200 205
 Pro Met Pro Leu Tyr Ser Leu Ser Leu Glu Glu Glu Val Glu Ala
 210 215 220
 Pro Val Leu Lys Leu Thr Ser Gly Asp Ser Gly Phe His Pro Glu Thr
 225 230 235 240
 Thr Asp Gln Val Leu Gln Glu Lys Lys Met Ala Leu Leu Thr Leu Leu
 245 250 255
 Cys Ser Ala Leu Ala Ser Asn Val Asn Val Lys Asp Ala Ser Asp Leu
 260 265 270
 Thr Arg Ala Ser Ile Leu Glu Val Cys Ser Ala Leu Ala Ser Leu Glu
 275 280 285
 Pro Glu Phe Ile Leu Lys Ala Ser Leu Tyr Ala Arg Gln Gln Leu Asn
 290 295 300
 Leu Arg Asp Ile Ala Asn Thr Val Leu Ala Val Ala Ala Leu Leu Pro
 305 310 315 320
 Ala Cys Arg Pro His Val Arg Arg Tyr Tyr Ser Ala Ile Val His Leu
 325 330 335
 Pro Ser Asp Trp Ile Gln Val Ala Glu Phe Tyr Gln Ser Leu Ala Glu
 340 345 350
 Gly Asp Glu Lys Lys Leu Val Ser Leu Pro Ala Cys Leu Arg Ala Ala
 355 360 365
 Met Thr Asp Lys Phe Ala Glu Phe Asp Glu Tyr Gln Leu Ala Lys Tyr
 370 375 380

Asn Pro Arg Lys His Arg Ser Lys Arg Arg Ser Arg Gln Pro Pro Arg
385 390 395 400

Pro Gln Lys Thr Glu Arg Pro Phe Ser Glu Arg Gly Lys Cys Phe Pro
405 410 415

Lys Ser Leu Trp Pro Leu Lys Asn Glu Gln Ile Thr Phe Glu Ala Ala
420 425 430

Tyr Asn Ala Met Pro Glu Lys Asn Arg Leu Pro Arg Phe Thr Leu Lys
435 440 445

Lys Leu Val Glu Tyr Leu His Ile His Lys Pro Ala Gln His Val Gln
450 455 460

Ala Leu Leu Gly Tyr Arg Tyr Pro Ala Thr Leu Glu Leu Phe Ser Arg
465 470 475 480

Ser His Leu Pro Gly Pro Trp Glu Ser Ser Arg Ala Gly Gln Arg Met
485 490 495

Lys Leu Arg Arg Pro Glu Thr Trp Glu Arg Glu Leu Ser Leu Arg Gly
500 505 510

Asn Lys Ala Ser Val Trp Glu Glu Leu Ile Asp Asn Gly Lys Leu Pro
515 520 525

Phe Met Ala Met Leu Arg Asn Leu Cys Asn Leu Leu Arg Thr Gly Ile
530 535 540

Ser Ala Arg His His Glu Leu Val Leu Gln Arg Leu Gln His Glu Lys
545 550 555 560

Ser Val Val His Ser Arg Gln Phe Pro Phe Arg Phe Leu Asn Ala His
565 570 575

Asp Ser Ile Asp Lys Leu Glu Ala Gln Leu Arg Ser Lys Ala Ser Pro
580 585 590

Phe Pro Ser Asn Thr Thr Leu Met Lys Arg Ile Met Ile Arg Asn Ser
595 600 605

Lys Lys Asn Arg Arg Pro Ala Ser Arg Lys His Leu Cys Thr Leu Thr
610 615 620

Arg Arg Gln Leu Arg Ala Ala Met Thr Ile Pro Val Met Tyr Glu Gln
625 630 635 640

Leu Lys Arg Glu Lys Leu Arg Leu His Lys Ala Arg Gln Trp Asn Cys
645 650 655

Asp Val Glu Leu Leu Glu Arg Tyr Arg Gln Ala Leu Glu Thr Ala Val
660 665 670

Asn Leu Ser Val Lys His Asn Leu Ser Pro Met Pro Gly Arg Thr Leu
675 680 685

Leu Val Tyr Leu Thr Asp Ala Asn Ala Asp Arg Leu Cys Pro Lys Ser
690 695 700

His Ser Gln Gly Pro Pro Leu Asn Tyr Val Leu Leu Leu Ile Gly Met
705 710 715 720

Met Val Ala Arg Ala Glu Gln Val Thr Val Cys Leu Cys Gly Gly Gly
 725 730 735
 Phe Val Lys Thr Pro Val Leu Thr Ala Asp Glu Gly Ile Leu Lys Thr
 740 745 750
 Ala Ile Lys Leu Gln Ala Gln Val Gln Glu Leu Glu Gly Asn Asp Glu
 755 760 765
 Trp Pro Leu Asp Thr Phe Gly Lys Tyr Leu Leu Ser Leu Ala Val Gln
 770 775 780
 Arg Thr Pro Ile Asp Arg Val Ile Leu Phe Gly Gln Arg Met Asp Thr
 785 790 795 800
 Glu Leu Leu Lys Val Ala Lys Gln Ile Ile Trp Gln His Val Asn Ser
 805 810 815
 Lys Cys Leu Phe Val Gly Val Leu Leu Gln Lys Thr Gln Tyr Ile Ser
 820 825 830
 Pro Asn Leu Asn Pro Asn Asp Val Thr Leu Ser Gly Cys Thr Asp Gly
 835 840 845
 Ile Leu Lys Phe Ile Ala Glu His Gly Ala Ser Arg Leu Leu Glu His
 850 855 860
 Val Gly Gln Leu Asp Lys Leu Phe Lys Ile Pro Pro Pro Pro Gly Lys
 865 870 875 880
 Thr Gln Ala Pro Ser Leu Arg Pro Leu Glu Glu Asn Ile Pro Gly Pro
 885 890 895
 Leu Gly Pro Ile Ser Gln His Gly Trp Arg Asn Ile Arg Leu Phe Ile
 900 905 910
 Ser Ser Thr Phe Arg Asp Met His Gly Glu Arg Asp Leu Leu Met Arg
 915 920 925
 Ser Val Leu Pro Ala Leu Gln Ala Arg Val Phe Pro His Arg Ile Ser
 930 935 940
 Leu His Ala Ile Asp Leu Arg Trp Gly Ile Thr Glu Glu Glu Thr Arg
 945 950 955 960
 Arg Asn Arg Gln Leu Glu Val Cys Leu Gly Glu Val Glu Asn Ser Gln
 965 970 975
 Leu Phe Val Gly Ile Leu Gly Ser Arg Tyr Gly Tyr Ile Pro Pro Ser
 980 985 990
 Tyr Asp Leu Pro Asp His Pro His Phe His Trp Thr His Glu Tyr Pro
 995 1000 1005
 Ser Gly Arg Ser Val Thr Glu Met Glu Val Met Gln Phe Leu Asn Arg
 1010 1015 1020
 Gly Gln Arg Ser Gln Pro Ser Ala Gln Ala Leu Ile Tyr Phe Arg Asp
 1025 1030 1035 1040
 Pro Asp Phe Leu Ser Ser Val Pro Asp Ala Trp Lys Pro Asp Phe Ile
 1045 1050 1055

Ser Glu Ser Glu Glu Ala Ala His Arg Val Ser Glu Leu Lys Arg Tyr
1060 1065 1070

Leu His Glu Gln Lys Glu Val Thr Cys Arg Ser Tyr Ser Cys Glu Trp
1075 1080 1085

Gly Gly Val Ala Ala Gly Arg Pro Tyr Thr Gly Gly Leu Glu Glu Phe
1090 1095 1100

Gly Gln Leu Val Leu Gln Asp Val Trp Ser Met Ile Gln Lys Gln His
1105 1110 1115 1120

Leu Gln Pro Gly Ala Gln Leu Glu Gln Pro Thr Ser Ile Ser Glu Asp
1125 1130 1135

Asp Leu Ile Gln Thr Ser Phe Gln Gln Leu Lys Thr Pro Thr Ser Pro
1140 1145 1150

Ala Arg Pro Arg Leu Leu Gln Asp Thr Val Gln Gln Leu Leu Leu Pro
1155 1160 1165

His Gly Arg Leu Ser Leu Val Thr Gly Gln Ala Gly Gln Gly Lys Thr
1170 1175 1180

Ala Phe Leu Ala Ser Leu Val Ser Ala Leu Lys Val Pro Asp Gln Pro
1185 1190 1195 1200

Asn Glu Pro Pro Phe Val Phe Phe His Phe Ala Ala Ala Arg Pro Asp
1205 1210 1215

Gln Cys Leu Ala Leu Asn Leu Leu Arg Arg Leu Cys Thr His Leu Arg
1220 1225 1230

Gln Lys Leu Gly Glu Leu Ser Ala Leu Pro Ser Thr Tyr Arg Gly Leu
1235 1240 1245

Val Trp Glu Leu Gln Gln Lys Leu Leu Leu Lys Phe Ala Gln Ser Leu
1250 1255 1260

Gln Pro Ala Gln Thr Leu Val Leu Ile Ile Asp Gly Ala Asp Lys Leu
1265 1270 1275 1280

Val Asp Arg Asn Gly Gln Leu Ile Ser Asp Trp Ile Pro Lys Ser Leu
1285 1290 1295

Pro Arg Arg Val His Leu Val Leu Ser Val Ser Ser Asp Ser Gly Leu
1300 1305 1310

Gly Glu Thr Leu Gln Gln Ser Gln Gly Ala Tyr Val Val Ala Leu Gly
1315 1320 1325

Ser Leu Val Pro Ser Ser Arg Ala Gln Leu Val Arg Glu Glu Leu Ala
1330 1335 1340

Leu Tyr Gly Lys Arg Leu Glu Glu Ser Pro Phe Asn Asn Gln Met Arg
1345 1350 1355 1360

Leu Leu Leu Ala Lys Gln Gly Ser Ser Leu Pro Leu Tyr Leu His Leu
1365 1370 1375

Val Thr Asp Tyr Leu Arg Leu Phe Thr Leu Tyr Glu Gln Val Ser Glu
1380 1385 1390

Arg Leu Arg Thr Leu Pro Ala Thr Leu Pro Leu Leu Leu Gln His Ile
1395 1400 1405

Leu Ser Thr Leu Glu Gln Glu His Gly His Asp Val Leu Pro Gln Ala
1410 1415 1420

Leu Thr Ala Leu Glu Val Thr Arg Ser Gly Leu Thr Val Asp Gln Leu
1425 1430 1435 1440

His Ala Ile Leu Ser Thr Trp Leu Ile Leu Pro Lys Glu Thr Lys Ser
1445 1450 1455

Trp Glu Glu Val Leu Ala Ala Ser His Ser Gly Asn Pro Phe Pro Leu
1460 1465 1470

Cys Pro Phe Ala Tyr Leu Val Gln Ser Leu Arg Ser Leu Leu Gly Glu
1475 1480 1485

Gly Pro Val Glu Arg Pro Gly Ala Arg Leu Cys Leu Ser Asp Gly Pro
1490 1495 1500

Leu Arg Thr Thr Ile Lys Arg Arg Tyr Gly Lys Arg Leu Gly Leu Glu
1505 1510 1515 1520

Lys Thr Ala His Val Leu Ile Ala Ala His Leu Trp Lys Thr Cys Asp
1525 1530 1535

Pro Asp Ala Ser Gly Thr Phe Arg Ser Cys Pro Pro Glu Ala Leu Lys
1540 1545 1550

Asp Leu Pro Tyr His Leu Leu Gln Ser Gly Asn His Gly Leu Leu Ala
1555 1560 1565

Glu Phe Leu Thr Asn Leu His Val Val Ala Ala Tyr Leu Glu Val Gly
1570 1575 1580

Leu Val Pro Asp Leu Leu Glu Ala His Val Leu Tyr Ala Ser Ser Lys
1585 1590 1595 1600

Pro Glu Ala Asn Gln Lys Leu Pro Ala Ala Asp Val Ala Val Phe His
1605 1610 1615

Thr Phe Leu Arg Gln Gln Ala Ser Leu Leu Thr Gln Tyr Pro Leu Leu
1620 1625 1630

Leu Leu Gln Gln Ala Ala Ser Gln Pro Glu Glu Ser Pro Val Cys Cys
1635 1640 1645

Gln Ala Pro Leu Leu Thr Gln Arg Trp His Asp Gln Phe Thr Leu Lys
1650 1655 1660

Trp Ile Asn Lys Pro Gln Thr Leu Lys Gly Gln Gln Ser Leu Ser Leu
1665 1670 1675 1680

Thr Met Ser Ser Ser Pro Thr Ala Val Ala Phe Ser Pro Asn Gly Gln
1685 1690 1695

Arg Ala Ala Val Gly Thr Ala Ser Gly Thr Ile Tyr Leu Leu Asn Leu
1700 1705 1710

Lys Thr Trp Gln Glu Glu Lys Ala Val Val Ser Gly Cys Asp Gly Ile
1715 1720 1725

Ser Ser Phe Ala Phe Leu Ser Asp Thr Ala Leu Phe Leu Thr Thr Phe
1730 1735 1740

Asp Gly His Leu Glu Leu Trp Asp Leu Gln His Gly Cys Trp Val Phe
1745 1750 1755 1760

Gln Thr Lys Ala His Gln Tyr Gln Ile Thr Gly Cys Cys Leu Ser Pro
1765 1770 1775

Asp Arg Arg Leu Leu Ala Thr Val Cys Leu Gly Gly Tyr Leu Lys Leu
1780 1785 1790

Trp Asp Thr Val Arg Gly Gln Leu Ala Phe Gln Tyr Thr His Pro Lys
1795 1800 1805

Ser Leu Asn Cys Val Ala Phe His Pro Glu Gly Gln Val Val Ala Thr
1810 1815 1820

Gly Ser Trp Ala Gly Ser Ile Thr Phe Phe Gln Ala Asp Gly Leu Lys
1825 1830 1835 1840

Val Thr Lys Glu Leu Gly Ala Pro Gly Pro Ser Val Cys Ser Leu Ala
1845 1850 1855

Phe Asn Lys Pro Gly Lys Ile Val Ala Val Gly Arg Ile Asp Gly Thr
1860 1865 1870

Val Glu Leu Trp Ala Trp Gln Glu Gly Ala Arg Leu Ala Ala Phe Pro
1875 1880 1885

Ala Gln Cys Gly Cys Val Ser Ala Val Leu Phe Leu His Ala Gly Asp
1890 1895 1900

Arg Phe Leu Thr Ala Gly Glu Asp Gly Lys Ala Gln Leu Trp Ser Gly
1905 1910 1915 1920

Phe Leu Gly Arg Pro Arg Gly Cys Leu Gly Ser Leu Pro Leu Ser Pro
1925 1930 1935

Ala Leu Ser Val Ala Leu Asn Pro Asp Gly Asp Gln Val Ala Val Gly
1940 1945 1950

Tyr Arg Glu Asp Gly Ile Asn Ile Tyr Lys Ile Ser Ser Gly Ser Gln
1955 1960 1965

Gly Pro Gln His Gln Glu Leu Asn Val Ala Val Ser Ala Leu Val Trp
1970 1975 1980

Leu Ser Pro Ser Val Leu Val Ser Gly Ala Glu Asp Gly Ser Leu His
1985 1990 1995 2000

Gly Trp Met Phe Lys Gly Asp Ser Leu His Ser Leu Trp Leu Leu Ser
2005 2010 2015

Arg Tyr Gln Lys Pro Val Leu Gly Leu Ala Ala Ser Arg Glu Leu Met
2020 2025 2030

Ala Ala Ala Ser Glu Asp Phe Thr Val Arg Leu Trp Pro Arg Gln Leu
2035 2040 2045

Leu Thr Gln Pro His Val His Ala Val Glu Leu Pro Cys Cys Ala Glu
2050 2055 2060

Leu Arg Gly His Glu Gly Pro Val Cys Cys Cys Ser Phe Ser Pro Asp
2065 2070 2075 2080

Gly Gly Ile Leu Ala Thr Ala Gly Arg Asp Arg Asn Leu Leu Cys Trp
2085 2090 2095

Asp Met Lys Ile Ala Gln Ala Pro Leu Leu Ile His Thr Phe Ser Ser
2100 2105 2110

Cys His Arg Asp Trp Ile Thr Gly Cys Ala Trp Thr Lys Asp Asn Ile
2115 2120 2125

Leu Val Ser Cys Ser Ser Asp Gly Ser Val Gly Leu Trp Asn Pro Glu
2130 2135 2140

Ala Gly Gln Gln Leu Gly Gln Phe Ser Gly His Gln Ser Ala Val Ser
2145 2150 2155 2160

Ala Val Val Ala Val Glu Glu His Ile Val Ser Val Ser Arg Asp Gly
2165 2170 2175

Thr Leu Lys Val Trp Asp His Gln Gly Val Glu Leu Thr Ser Ile Pro
2180 2185 2190

Ala His Ser Gly Pro Ile Ser Gln Cys Ala Ala Ala Leu Glu Pro Arg
2195 2200 2205

Pro Gly Gly Gln Pro Gly Ser Glu Leu Leu Val Val Thr Val Gly Leu
2210 2215 2220

Asp Gly Ala Thr Lys Leu Trp His Pro Leu Leu Val Cys Gln Ile Arg
2225 2230 2235 2240

Thr Leu Gln Gly His Ser Gly Pro Val Thr Ala Ala Ala Ala Ser Glu
2245 2250 2255

Ala Ser Gly Leu Leu Leu Thr Ser Asp Asp Ser Ser Val Gln Leu Trp
2260 2265 2270

Gln Ile Pro Lys Glu Ala Asp Asp Ser Tyr Lys Pro Arg Ser Ser Val
2275 2280 2285

Ala Ile Thr Ala Val Ala Trp Ala Pro Asp Gly Ser Met Val Val Ser
2290 2295 2300

Gly Asn Glu Ala Gly Glu Leu Thr Leu Trp Gln Gln Ala Lys Ala Val
2305 2310 2315 2320

Ala Thr Ala Gln Ala Pro Gly Arg Val Ser His Leu Ile Trp Tyr Ser
2325 2330 2335

Ala Asn Ser Phe Phe Val Leu Ser Ala Asn Glu Asn Val Ser Glu Trp
2340 2345 2350

Gln Val Gly Leu Arg Lys Gly Ser Thr Ser Thr Ser Ser Leu His
2355 2360 2365

Leu Lys Arg Val Leu Gln Glu Asp Trp Gly Val Leu Thr Gly Leu Gly
2370 2375 2380

Leu Ala Pro Asp Gly Gln Ser Leu Ile Leu Met Lys Glu Asp Val Glu
2385 2390 2395 2400

Leu Leu Glu Met Lys Pro Gly Ser Ile Pro Ser Ser Ile Cys Arg Arg
 2405 2410 2415
 Tyr Gly Val His Ser Ser Ile Leu Cys Thr Ser Lys Glu Tyr Gly Leu
 2420 2425 2430
 Phe Tyr Leu Gln Gln Gly Asp Ser Gly Leu Leu Ser Ile Leu Glu Gln
 2435 2440 2445
 Lys Glu Ser Gly Glu Phe Glu Glu Ile Leu Asp Phe Asn Leu Asn Leu
 2450 2455 2460
 Asn Asn Pro Asn Gly Ser Pro Val Ser Ile Thr Gln Ala Lys Pro Glu
 2465 2470 2475 2480
 Ser Glu Ser Ser Leu Leu Cys Ala Thr Ser Asp Gly Met Leu Trp Asn
 2485 2490 2495
 Leu Ser Glu Cys Thr Ser Glu Gly Glu Trp Ile Val Asp Asn Ile Trp
 2500 2505 2510
 Gln Lys Lys Ala Lys Lys Pro Lys Thr Gln Thr Leu Glu Thr Glu Leu
 2515 2520 2525
 Ser Pro His Ser Glu Leu Asp Phe Ser Ile Asp Cys Trp Ile Asp Pro
 2530 2535 2540
 Thr Asn Leu Lys Ala Gln Gln Cys Lys Ile His Leu Gly Ser Val
 2545 2550 2555 2560
 Thr Ala Leu His Val Leu Pro Gly Leu Leu Val Thr Ala Ser Lys Asp
 2565 2570 2575
 Arg Asp Val Lys Leu Trp Glu Arg Pro Ser Met Gln Leu Leu Gly Leu
 2580 2585 2590
 Phe Arg Cys Glu Gly Pro Val Ser Cys Leu Glu Pro Trp Met Glu Pro
 2595 2600 2605
 Ser Ser Pro Leu Gln Leu Ala Val Gly Asp Thr Gln Gly Asn Leu Tyr
 2610 2615 2620
 Phe Leu Ser Trp Glu
 2625

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCTCTGCGGC CGCTACANNNNNNNT

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGAGACGCCG GCGA

14

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TCGACCCACG CGTCCG

16

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGTGCGCAG GC

12

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGTAAAACGA CGGCCAGT

18

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CAGGAAACAG CTATGACC

18

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CAATTAACCC TCACTAAAG

19

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 154 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGTACCGCCA GCCGAGCCAC ATCGCTCAGA CACCATGATC GCAAATGTGA ATATTGCTCA

60

GGAACAAAAG CTTATTTCTG AAGAAGACTT GGCTCAGGAA CAAAAGCTTA TTTCTGAAGA

120

AGACTTGGCT CAGCAGAGTG GCGGAGGACT CGAG

154

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2848 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CACGCGTCCG	GGCAGCGCTG	CGTCCTGCTG	CGCACGTGGG	AAGCCCTGGC	CCC GGCC ACC	60
CCCGCGATGC	CGCGCGCTCC	CCGCTGCCGA	GCCGTGCGCT	CCCTGCTGCG	CAGCCACTAC	120
CGCGAGGTGC	TGCCGCTGGC	CACGTTCGTG	CGGCGCCTGG	GGCCCCAGGG	CTGGCGGCTG	180
GTGCAGCGCG	GGGACCCGGC	GGCTTCCGC	GCGCTGGTGG	CCCAGTGCCT	GGTGTGCGTG	240
CCCTGGGACG	CACGGCCGCC	CCCCGCCGCC	CCCTCCTTCC	GCCAGGGTGT	CTGCCTGAAG	300
GAGCTGGTGG	CCCGAGTGCT	GCAGAGGCTG	TGCGAGCGCG	GCGCGAAAGAA	CGTGCTGGCC	360
TCGGGCTTCG	CGCTGCTGGA	CGGGGCCCCG	GGGGGGCCCC	CCGAGGCCTT	CACCACCAGC	420
GTGCGCAGCT	ACCTGCCAA	CACGGTGACC	GACGCACTGC	GGGGGAGCGG	GGCGTGGGGG	480
CTGCTGCTGC	GCCGGTGGG	CGACGACGTG	CTGGTTCA	TGCTGGCACG	CTGCGCGCTC	540
TTTGTGCTGG	TGGCTCCCAG	CTGCGCCTAC	CAGGTGTGCG	GGCCGCCGCT	GTACCAGCTC	600
GGCGCTGCCA	CTCAGGCCCG	GCCCCCGCCA	CACGCTAGTG	GACCCCGAAG	CGTCTGGGA	660
TGCGAACGGG	CCTGGAACCA	TAGCGTCAGG	GAGGCCGGGG	TCCCCCTGGG	CCTGCCAGCC	720
CCGGGTGCCA	GGAGGCGCGG	GGGCAGTGCC	AGCCGAAGTC	TGCCGTTGCC	CAAGAGGCC	780
AGGCGTGGCG	CTGCCCCTGA	GCCGGAGCGG	ACGCCCGTTG	GGCAGGGTC	CTGGGCCAC	840
CCGGGCAGGA	CGCGTGGACC	GAGTGACCGT	GGTTTCTGTG	TGGTGTCA	TGCCAGACCC	900
GCCGAAGAAG	CCACCTCTTT	GGAGGGTGCG	CTCTCTGGCA	CGCGCCACTC	CCACCCATCC	960
GTGGGCCGCC	AGCACCAACGC	GGGCCCCCA	TCCACATCGC	GGCCACCACG	TCCCTGGGAC	1020
ACGCCTTGTG	CCCCGGTGT	CGCCGAGACC	AAGCACTTCC	TCTACTCCTC	AGGCGACAAG	1080
GAGCAGCTGC	GGCCCTCCTT	CCTACTCAGC	TCTCTGAGGC	CCAGCCTGAC	TGGCGCTCGG	1140
AGGCTCGTGG	AGACCATCTT	TCTGGGTTC	AGGCCCTGG	TGCCAGGGAC	TCCCCGCAGG	1200
TTGCCCGGCC	TGCCCGAGCG	CTACTGGCAA	ATGCCGCC	TGTTTCTGG	GCTGCTTGGG	1260
AACCACGCGC	AGTGCCCTA	CGGGGTGCTC	CTCAAGACGC	ACTGCCGCT	GCGAGCTGCG	1320
GTCACCCAG	CAGCCGGTGT	CTGTGCCCGG	GAGAAGCCCC	AGGGCTCTGT	GGCGGCC	1380
GAGGAGGAGG	ACACAGACCC	CCGTCGCCTG	GTGCAGCTGC	TCCGCCAGCA	CAGCAGCCCC	1440
TGGCAGGTGT	ACGGCTTCGT	GGGGGCCTGC	CTGCGCCGGC	TGGTCCCCC	AGGCCTCTGG	1500
GGCTCCAGGC	ACAACGAACG	CCGCTTCC	AGGAACACCA	AGAAGTTCAT	CTCCCTGGGG	1560
AAGCATGCCA	AGCTCTCGCT	GCAGGAGCTG	ACGTGGAAGA	TGAGCGTGCG	GGACTGCGCT	1620
TGGCTCGCA	GGAGGCCAGG	GGTTGGCTGT	GTTCCGGCCG	CAGAGCACCG	TCTGCGTGAG	1680
GAGATCCTGG	CCAAGTTCC	GCACTGGCTG	ATGAGTGTGT	ACGTCGTCGA	GCTGCTCAGG	1740
TCTTTCTTTT	ATGTCACGGA	GACCACGTTT	CAAAAGAAC	GGCTCTTTT	CTACCGGAAG	1800

AGTGTCTGGA	GCAAGTTGCA	AAGCATTGGA	ATCAGACAGC	ACTTGAAGAG	GGTGCAGCTG	1860
CGGGAGCTGT	CGGAAGCAGA	GGTCAGGCAG	CATCGGGAAG	CCAGGCCCGC	CCTGCTGACG	1920
TCCAGACTCC	GCTTCATCCC	CAAGCCTGAC	GGGCTGCGGC	CGATTGTGAA	CATGGACTAC	1980
GTCGTGGGAG	CCAGAACGTT	CCGCAGAGAA	AAGAGGGCCG	AGCGTCTCAC	CTCGAGGGTG	2040
AAGGCACTGT	TCAGCGTGCT	CAAATACCGAG	CGGGCGCGGC	GCCCCGGCCT	CCTGGGCGCC	2100
TCTGTGCTGG	GCCTGGACGA	TATCCACAGG	GCCTGGCGCA	CCTTCGTGCT	GCGTGTGCGG	2160
GCCCAGGACC	CGCCGCCTGA	GCTGTACTTT	GTCAAGGTGG	ATGTGACGGG	CGCGTACGAC	2220
ACCATCCCCC	AGGACAGGCT	CACGGAGGTC	ATCGCCAGCA	TCATCAAACC	CCAGAACACG	2280
TACTGCGTGC	GTCGGTATGC	CGTGGTCCAG	AAGGCCGCC	ATGGGCACGT	CCGCAAGGCC	2340
TTCAAGAGCC	ACGTCTCTAC	CTTGACAGAC	CTCCAGCCGT	ACATGCGACA	GTTCGTGGCT	2400
CACCTGCAGG	AGACCAGCCC	GCTGAGGGAT	GCCGTCGTCA	TCGAGCAGAG	CTCCTCCCTG	2460
AATGAGGCCA	GCAGTGGCCT	CTTCGACGTC	TTCCCTACGCT	TCATGTGCCA	CCACGCCGTG	2520
CGCATCAGGG	GCAAGTCCTA	CGTCCAGTGC	CAGGGGATCC	CCGAGGGCTC	CATCCTCTCC	2580
ACGCTGCTCT	GCAGCCTGTG	CTACGGCGAC	ATGGAGAACAA	AGCTGTTTGC	GGGGATTCCG	2640
CGGGACGGGC	TGCTCCTGCG	TTTGGTGGAT	GATTCTTGT	TGGTGACACC	TCACCTCACC	2700
CACGCGAAAA	CCTTCCTCAG	GACCCTGGTC	CGAGGTGTCC	CTGAGTATGG	CTGCGTGGTG	2760
AACTTGCAGA	AGACAGTGGT	GAACCTCCCT	GTAGAAGACG	AGGCCCTGGG	TGGCACGGCT	2820
TTTGTTCAGA	TGCCGGCCCA	CGGCCTAT				2848

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 949 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His	Ala	Ser	Gly	Gln	Arg	Cys	Val	Leu	Leu	Arg	Thr	Trp	Glu	Ala	Leu
1				5					10				15		
Ala	Pro	Ala	Thr	Pro	Ala	Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val
	20					25						30			
Arg	Ser	Leu	Leu	Arg	Ser	His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr
		35					40					45			
Phe	Val	Arg	Arg	Leu	Gly	Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly
	50				55					60					
Asp	Pro	Ala	Ala	Phe	Arg	Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val
	65				70					75			80		

Pro Trp Asp Ala Arg Pro Pro Pro Ala Ala Pro Ser Phe Arg Gln Val
85 90 95

Ser Cys Leu Lys Glu Leu Val Ala Arg Val Leu Gln Arg Leu Cys Glu
100 105 110

Arg Gly Ala Lys Asn Val Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly
115 120 125

Ala Arg Gly Gly Pro Pro Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr
130 135 140

Leu Pro Asn Thr Val Thr Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly
145 150 155 160

Leu Leu Leu Arg Arg Val Gly Asp Asp Val Leu Val His Leu Leu Ala
165 170 175

Arg Cys Ala Leu Phe Val Leu Val Ala Pro Ser Cys Ala Tyr Gln Val
180 185 190

Cys Gly Pro Pro Leu Tyr Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro
195 200 205

Pro Pro His Ala Ser Gly Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala
210 215 220

Trp Asn His Ser Val Arg Glu Ala Gly Val Pro Leu Gly Leu Pro Ala
225 230 235 240

Pro Gly Ala Arg Arg Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu
245 250 255

Pro Lys Arg Pro Arg Arg Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro
260 265 270

Val Gly Gln Gly Ser Trp Ala His Pro Gly Arg Thr Arg Gly Pro Ser
275 280 285

Asp Arg Gly Phe Cys Val Val Ser Pro Ala Arg Pro Ala Glu Glu Ala
290 295 300

Thr Ser Leu Glu Gly Ala Leu Ser Gly Thr Arg His Ser His Pro Ser
305 310 315 320

Val Gly Arg Gln His His Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro
325 330 335

Arg Pro Trp Asp Thr Pro Cys Pro Pro Val Tyr Ala Glu Thr Lys His
340 345 350

Phe Leu Tyr Ser Ser Gly Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu
355 360 365

Leu Ser Ser Leu Arg Pro Ser Leu Thr Gly Ala Arg Arg Leu Val Glu
370 375 380

Thr Ile Phe Leu Gly Ser Arg Pro Trp Met Pro Gly Thr Pro Arg Arg
385 390 395 400

Leu Pro Arg Leu Pro Gln Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu
405 410 415

Glu Leu Leu Gly Asn His Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys
420 425 430

Thr His Cys Pro Leu Arg Ala Ala Val Thr Pro Ala Ala Gly Val Cys
435 440 445

Ala Arg Glu Lys Pro Gln Gly Ser Val Ala Ala Pro Glu Glu Glu Asp
450 455 460

Thr Asp Pro Arg Arg Leu Val Gln Leu Leu Arg Gln His Ser Ser Pro
465 470 475 480

Trp Gln Val Tyr Gly Phe Val Arg Ala Cys Leu Arg Arg Leu Val Pro
485 490 495

Pro Gly Leu Trp Gly Ser Arg His Asn Glu Arg Arg Phe Leu Arg Asn
500 505 510

Thr Lys Lys Phe Ile Ser Leu Gly Lys His Ala Lys Leu Ser Leu Gln
515 520 525

Glu Leu Thr Trp Lys Met Ser Val Arg Asp Cys Ala Trp Leu Arg Arg
530 535 540

Ser Pro Gly Val Gly Cys Val Pro Ala Ala Glu His Arg Leu Arg Glu
545 550 555 560

Glu Ile Leu Ala Lys Phe Leu His Trp Leu Met Ser Val Tyr Val Val
565 570 575

Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys
580 585 590

Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val Trp Ser Lys Leu Gln Ser
595 600 605

Ile Gly Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg Glu Leu Ser
610 615 620

Glu Ala Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu Leu Thr
625 630 635 640

Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val
645 650 655

Asn Met Asp Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg
660 665 670

Ala Glu Arg Leu Thr Ser Arg Val Lys Ala Leu Phe Ser Val Leu Asn
675 680 685

Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser Val Leu Gly
690 695 700

Leu Asp Asp Ile His Arg Ala Trp Arg Thr Phe Val Leu Arg Val Arg
705 710 715 720

Ala Gln Asp Pro Pro Pro Glu Leu Tyr Phe Val Lys Val Asp Val Thr
725 730 735

Gly Ala Tyr Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala
740 745 750

Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val
 755 760 765
 Val Gln Lys Ala Ala His Gly His Val Arg Lys Ala Phe Lys Ser His
 770 775 780
 Val Ser Thr Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln Phe Val Ala
 785 790 795 800
 His Leu Gln Glu Thr Ser Pro Leu Arg Asp Ala Val Val Ile Glu Gln
 805 810 815
 Ser Ser Ser Leu Asn Glu Ala Ser Ser Gly Leu Phe Asp Val Phe Leu
 820 825 830
 Arg Phe Met Cys His His Ala Val Arg Ile Arg Gly Lys Ser Tyr Val
 835 840 845
 Gln Cys Gln Gly Ile Pro Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys
 850 855 860
 Ser Leu Cys Tyr Gly Asp Met Glu Asn Lys Leu Phe Ala Gly Ile Arg
 865 870 875 880
 Arg Asp Gly Leu Leu Leu Arg Leu Val Asp Asp Phe Leu Leu Val Thr
 885 890 895
 Pro His Leu Thr His Ala Lys Thr Phe Leu Arg Thr Leu Val Arg Gly
 900 905 910
 Val Pro Glu Tyr Gly Cys Val Val Asn Leu Arg Lys Thr Val Val Asn
 915 920 925
 Phe Pro Val Glu Asp Glu Ala Leu Gly Gly Thr Ala Phe Val Gln Met
 930 935 940
 Pro Ala His Gly Leu
 945

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCAAGTTCCCT GCACTGGCTG AT

22

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCTCGTAGTT GAGCACGCTG AA

22

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Phe Phe Tyr Val Thr Glu
 1 5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 949 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TCCCCTGGTG CGGCCTGCTG CTGGATACCC GGACCCCTGGA GGTGCAGAGC GACTACTCCA	60
GCTATGCCCG GACCTCCATC AGAGCCAGTC TCACCTTCAA CCGCGGCTTC AAGGCTGGGA	120
GGAACATGCG TCGCAAACTC TTTGGGGTCT TGCGGCTGAA GTGTACAGC CTGTTCTGG	180
ATTTGCAGGT GAACAGCCTC CAGACGGTGT GCACCAACAT CTACAAGATC CTCCTGCTGC	240
AGGCGTACAG GTTTCACGCA TGTGTGCTGC AGCTCCCATT TCATCAGCAA GTTTGGAAGA	300
ACCCCCACATT TTTCCTGCGC GTCATCTCTG ACACGGCCTC CCTCTGCTAC TCCATCCTGA	360
AAGCCAAGAA CGCAGGGATG TCGCTGGGG CCAAGGGCGC CGCCGGCCCT CTGCCCTCCG	420
AGGCCGTGCA GTGGCTGTGC CACCAAGCAT TCCTGCTCAA GCTGACTCGA CACCGTGTCA	480
CCTACGTGCC ACTCCTGGGG TCACTCAGGA CAGCCCAGAC GCAGCTGAGT CGGAAGCTCC	540
CGGGGACGAC GCTGACTGCC CTGGAGGCCG CAGCCAACCC GGCACTGCC TCAGACTTCA	600
AGACCATCCT GGACTGATGG CCACCCGCC ACAGCCAGGC CGAGAGCAGA CACCAGCAGC	660
CCTGTCACGC CGGGCTCTAC GTCCCAGGGA GGGAGGGCG GCCCACACCC AGGCCCGCAC	720
CGCTGGGAGT CTGAGGCCTG AGTGAGTGTT TGGCCGAGGC CTGCATGTCC GGCTGAAGGC	780

TGAGTGTCCG GCTGAGGCCT GAGCGAGTGT CCAGCCAAGG GCTGAGTGTC CAGCACACCT	840
GCCGTCTTCA CTTCCCCACA GGCTGGCGCT CGGCTCCACC CCAGGGCCAG CTTTCCTCA	900
CCAGGAGCCC GGCTTCCACT CCCACATAG GAATAGTCCA TCCCCTGAT	949

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3798 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCACCGTCC GGGCAGCGCT GCGTCCTGCT GCGCACGTGG GAAGCCCTGG CCCCAGGCCAC	60
CCCCCGCATG CCGCGCGCTC CCCGCTGCCG AGCCGTGCGC TCCCTGCTGC GCAGCCACTA	120
CCCGGAGGTG CTGCCGCTGG CCACGTTCGT GCGGCCGCTG GGGCCCCAGG GCTGGCGGCT	180
GGTGCAGCGC GGGGACCCGG CGGCTTCCG CGCGCTGGTG GCCCAGTGCC TGGTGTGCGT	240
GCCCTGGGAC GCACGGCCGC CCCCCGCCGC CCCCTCCTTC CGCCAGGTGT CCTGCCTGAA	300
GGAGCTGGTG GCCCGAGTGC TGCAGAGGCT GTGCGAGCGC GGCGCGAAGA ACGTGCTGGC	360
CTTCGGCTTC GCGCTGCTGG ACGGGGCCCG CGGGGGCCCC CCCGAGGCCT TCACCACCAAG	420
CGTGCAGCAGC TACCTGCCCA ACACGGTGAC CGACGCAGTG CGGGGGAGCG GGGCGTGGGG	480
GCTGCTGCTG CGCCCGTGG GCGACGACGT GCTGGTTCAC CTGCTGGCAC GCTGCGCGCT	540
CTTTGTGCTG GTGGCTCCA GCTGCGCTA CCAGGTGTGC GGGCCGCCGC TGTACCAAGCT	600
CGGCGCTGCC ACTCAGGCCCG GGCCCCCGCC ACACGCTAGT GGACCCCGAA GGCGTCTGGG	660
ATGCGAACGG GCCTTGAACC ATAGCGTCAG GGAGGGCGGG GTCCCCCTGG GCCTGCCAGC	720
CCCGGGTGGC AGGAGGCGCG GGGCAGTGC CAGCCGAAGT CTGCCGTTGC CCAAGAGGCC	780
CAGGGCGTGGC GCTGCCCTG AGCCGGAGCG GACGCCCGTT GGGCAGGGGT CCTGGGCCCA	840
CCCGGGCAGG ACGCGTGGAC CGAGTGACCG TGGTTCTGT GTGGTGTAC CTGCCAGACC	900
CGCCGAAGAA GCCACCTCTT TGGAGGGTGC GCTCTCTGGC ACGGCCACT CCCACCCATC	960
CGTGGGCCGC CAGCACCAAG CGGGCCCCCCC ATCCACATCG CGGCCACAC GTCCCTGGGA	1020
CACGCCCTGT CCCCCGGTGT ACGCCGAGAC CAAGCACTTC CTCTACTCCT CAGGCGACAA	1080
GGAGCAGCTG CGGCCCTCCT TCCTACTCAG CTCTCTGAGG CCCAGCCTGA CTGGCGCTCG	1140
GAGGCTCGTG GAGACCATCT TTCTGGGTTG CAGGCCCTGG ATGCCAGGGA CTCCCCGCAG	1200
GTTGCCCGGC CTGCCCGAGC GCTACTGGCA AATGCCGGCCC CTGTTCTGG AGCTGCTTGG	1260
GAACCACGGC CAGTGCCCT ACGGGGTGCT CCTCAAGACG CACTGCCGC TGCGAGCTGC	1320

GGTCACCCCA	GCAGCCGGTG	TCTGTGCCCG	GGAGAAAGCCC	CAGGGCTCTG	TGGCGGCC	1380
CGAGGAGGAG	GACACAGACC	CCCGTCGCCCT	GGTGCAGCTG	CTCCGCCAGC	ACAGCAGCCC	1440
CTGGCAGGTG	TACGGCTTCG	TGCGGGCCTG	CCTGCGCCGG	CTGGTGCCCC	CAGGCCTCTG	1500
GGGCTCCAGG	CACAACGAAC	GCCGCTTCCT	CAGGAACACC	AAGAACGTTCA	TCTCCCTGGG	1560
GAAGCATGCC	AAGCTCTCGC	TGCAGGAGCT	GACGTGGAAG	ATGAGCGTGC	GGGACTGCGC	1620
TTGGCTGCGC	AGGAGCCCAG	GGGTTGGCTG	TGTTCCGGCC	GCAGAGCACC	GTCTGCGTGA	1680
GGAGATCCTG	GCCAAGTTCC	TGCACTGGCT	GATGAGTGTG	TACGTCGTCG	AGCTGCTCAG	1740
GTCTTCTTT	TATGTCACGG	AGACCACGTT	TCAAAAGAAC	AGGCTCTTT	TCTACCGGAA	1800
GAGTGTCTGG	AGCAAGTTGC	AAAGCATTGG	AATCAGACAG	CACTGAAGA	GGGTGCAGCT	1860
GCGGGAGCTG	TCGGAAGCAG	AGGTCAAGGCA	GCATCGGGAA	GCCAGGCC	CCCTGCTGAC	1920
GTCCAGACTC	CGCTTCATCC	CCAAGCCTGA	CGGGCTGCGG	CCGATTGTGA	ACATGGACTA	1980
CGTCGTGGGA	GCCAGAACGT	TCCGCAGAGA	AAAGAGGGCC	GAGCGTCTCA	CCTCGAGGGT	2040
GAAGGCACTG	TTCAGCGTGC	TCAACTACGA	GCAGGGCGCGG	CGCCCCGGCC	TCCTGGCGC	2100
CTCTGTGCTG	GGCCTGGACG	ATATCCACAG	GGCCTGGCGC	ACCTTCGTGC	TGCGTGTGCG	2160
GGCCCAGGAC	CCGCCGCTG	AGCTGTACTT	TGTCAGGTG	GATGTGACGG	GCGCGTACGA	2220
CACCATCCCC	CAGGACAGGC	TCACGGAGGT	CATCGCCAGC	ATCATCAAAC	CCCAGAACAC	2280
GTACTGCGTG	CGTCGGTATG	CCGTGGTCCA	GAAGGCGGCC	CATGGGCACG	TCCGCAAGGC	2340
CTTCAAGAGC	CACGTCTCTA	CCTTGACAGA	CCTCCAGCCG	TACATGCGAC	AGTTCGTGGC	2400
TCACCTGCGAG	GAGACCAGCC	CGCTGAGGGG	TGCCGTCGTC	ATCGAGCAGA	GCTCCTCCCT	2460
GAATGAGGCC	AGCAGTGGCC	TCTTCGACGT	CTTCCTACGC	TTCATGTGCC	ACCACGCCGT	2520
GCGCATCAGG	GGCAAGTCCT	ACGTCCAGTG	CCAGGGGATC	CCGCAGGGCT	CCATCCTCTC	2580
CACGCTGCTC	TGCAGCCTGT	GCTACGGCGA	CATGGAGAAC	AAGCTTTTG	CGGGGATTG	2640
GCAGGACGGG	CTGCTCCTGC	GTGGTGGGA	TGATTTCTTG	TTGGTGACAC	CTCACCTCAC	2700
CCACCGAAA	ACCTCCTCA	GGACCTGGT	CCGAGGTGTC	CCTGAGTATG	GCTGCGTGGT	2760
GAACCTGCGG	AAGACAGTGG	TGAACCTCCC	TGTAGAACAC	GAGGCCCTGG	GTGGCACGGC	2820
TTTTGTTCAAG	ATGCCGGCCC	ACGGCCTATT	CCCCTGGTGC	GGCCTGCTGC	TGGATAACCG	2880
GACCCCTGGAG	GTGCAGAGCG	ACTACTCCAG	CTATGCCCGG	ACCTCCATCA	GAGCCAGTCT	2940
CACCTTCAAC	CGCGGCTTCA	AGGCTGGGAG	GAACATGCGT	CGCAAACCT	TTGGGGTCTT	3000
GCGGCTGAAG	TGTCACAGCC	TGTTTCTGGA	TTTGCAGGTG	AACAGCCTCC	AGACGGTGTG	3060
CACCAACATC	TACAAGATCC	TCCTGCTGCA	GGCGTACAGG	TTTCACGCAT	GTGTGCTGCA	3120
GCTCCCATTT	CATCAGCAAG	TTTGGAAAGAA	CCCCACATTT	TTCCCTGCGCG	TCATCTCTGA	3180

CACGGCCTCC	CTCTGCTACT	CCATCCTGAA	AGCCAAGAAC	GCAGGGATGT	CGCTGGGGC	3240
CAAGGGCGCC	GCCGCCCTC	TGCCCTCCGA	GGCCGTGCAG	TGGCTGTGCC	ACCAAGCATT	3300
CCTGCTCAAG	CTGACTCGAC	ACCGTGTAC	CTACGTGCCA	CTCCTGGGT	CACTCAGGAC	3360
AGCCCAGACG	CAGCTGAGTC	GGAAGCTCCC	GGGGACGACG	CTGACTGCC	TGGAGGCCG	3420
AGCCAACCCG	GCACTGCCCT	CAGACTTCAA	GACCATCCTG	GAUTGATGGC	CACCCGCCA	3480
CAGCCAGGCC	GAGAGCAGAC	ACCAGCAGCC	CTGTCACGCC	GGGCTCTACG	TCCCAGGGAG	3540
GGAGGGGCCG	CCCACACCCA	GGCCCGCACC	GCTGGGAGTC	TGAGGCCTGA	GTGAGTGT	3600
GGCCGAGGCC	TGCATGTCCG	GCTGAAGGCT	GAGTGTCCGG	CTGAGGCCTG	AGCGAGTGTC	3660
CAGCCAAGGG	CTGAGTGTCC	AGCACACCTG	CCGTCTTCAC	TTCCCCACAG	GCTGGCGCTC	3720
GGCTCCACCC	CAGGCCAGC	TTTCCTCAC	CAGGAGCCCG	GCTTCCACTC	CCCACATAGG	3780
AATA GTCCAT	CCCCTGAT					3798

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1154 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

His	Ala	Ser	Gly	Gln	Arg	Cys	Val	Leu	Leu	Arg	Thr	Trp	Glu	Ala	Leu
1				5					10					15	
Ala	Pro	Ala	Thr	Pro	Ala	Met	Pro	Arg	Ala	Pro	Arg	Cys	Arg	Ala	Val
	20					25				30					
Arg	Ser	Leu	Leu	Arg	Ser	His	Tyr	Arg	Glu	Val	Leu	Pro	Leu	Ala	Thr
	35					40					45				
Phe	Val	Arg	Arg	Leu	Gly	Pro	Gln	Gly	Trp	Arg	Leu	Val	Gln	Arg	Gly
	50				55				60						
Asp	Pro	Ala	Ala	Phe	Arg	Ala	Leu	Val	Ala	Gln	Cys	Leu	Val	Cys	Val
	65				70				75		80				
Pro	Trp	Asp	Ala	Arg	Pro	Pro	Ala	Ala	Pro	Ser	Phe	Arg	Gln	Val	
		85				90					95				
Ser	Cys	Leu	Lys	Glu	Leu	Val	Ala	Arg	Val	Leu	Gln	Arg	Leu	Cys	Glu
		100				105				110					
Arg	Gly	Ala	Lys	Asn	Val	Leu	Ala	Phe	Gly	Phe	Ala	Leu	Leu	Asp	Gly
		115				120					125				
Ala	Arg	Gly	Gly	Pro	Pro	Glu	Ala	Phe	Thr	Thr	Ser	Val	Arg	Ser	Tyr
	130				135					140					

Leu Pro Asn Thr Val Thr Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly
145 150 155 160

Leu Leu Leu Arg Arg Val Gly Asp Asp Val Leu Val His Leu Leu Ala
165 170 175

Arg Cys Ala Leu Phe Val Leu Val Ala Pro Ser Cys Ala Tyr Gln Val
180 185 190

Cys Gly Pro Pro Leu Tyr Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro
195 200 205

Pro Pro His Ala Ser Gly Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala
210 215 220

Trp Asn His Ser Val Arg Glu Ala Gly Val Pro Leu Gly Leu Pro Ala
225 230 235 240

Pro Gly Ala Arg Arg Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu
245 250 255

Pro Lys Arg Pro Arg Arg Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro
260 265 270

Val Gly Gln Gly Ser Trp Ala His Pro Gly Arg Thr Arg Gly Pro Ser
275 280 285

Asp Arg Gly Phe Cys Val Val Ser Pro Ala Arg Pro Ala Glu Glu Ala
290 295 300

Thr Ser Leu Glu Gly Ala Leu Ser Gly Thr Arg His Ser His Pro Ser
305 310 315 320

Val Gly Arg Gln His His Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro
325 330 335

Arg Pro Trp Asp Thr Pro Cys Pro Pro Val Tyr Ala Glu Thr Lys His
340 345 350

Phe Leu Tyr Ser Ser Gly Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu
355 360 365

Leu Ser Ser Leu Arg Pro Ser Leu Thr Gly Ala Arg Arg Leu Val Glu
370 375 380

Thr Ile Phe Leu Gly Ser Arg Pro Trp Met Pro Gly Thr Pro Arg Arg
385 390 395 400

Leu Pro Arg Leu Pro Gln Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu
405 410 415

Glu Leu Leu Gly Asn His Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys
420 425 430

Thr His Cys Pro Leu Arg Ala Ala Val Thr Pro Ala Ala Gly Val Cys
435 440 445

Ala Arg Glu Lys Pro Gln Gly Ser Val Ala Ala Pro Glu Glu Glu Asp
450 455 460

Thr Asp Pro Arg Arg Leu Val Gln Leu Leu Arg Gln His Ser Ser Pro
465 470 475 480

Trp Gln Val Tyr Gly Phe Val Arg Ala Cys Leu Arg Arg Leu Val Pro
485 490 495

Pro Gly Leu Trp Gly Ser Arg His Asn Glu Arg Arg Phe Leu Arg Asn
500 505 510

Thr Lys Lys Phe Ile Ser Leu Gly Lys His Ala Lys Leu Ser Leu Gln
515 520 525

Glu Leu Thr Trp Lys Met Ser Val Arg Asp Cys Ala Trp Leu Arg Arg
530 535 540

Ser Pro Gly Val Gly Cys Val Pro Ala Ala Glu His Arg Leu Arg Glu
545 550 555 560

Glu Ile Leu Ala Lys Phe Leu His Trp Leu Met Ser Val Tyr Val Val
565 570 575

Glu Leu Leu Arg Ser Phe Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys
580 585 590

Asn Arg Leu Phe Phe Tyr Arg Lys Ser Val Trp Ser Lys Leu Gln Ser
595 600 605

Ile Gly Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg Glu Leu Ser
610 615 620

Glu Ala Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu Leu Thr
625 630 635 640

Ser Arg Leu Arg Phe Ile Pro Lys Pro Asp Gly Leu Arg Pro Ile Val
645 650 655

Asn Met Asp Tyr Val Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg
660 665 670

Ala Glu Arg Leu Thr Ser Arg Val Lys Ala Leu Phe Ser Val Leu Asn
675 680 685

Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu Gly Ala Ser Val Leu Gly
690 695 700

Leu Asp Asp Ile His Arg Ala Trp Arg Thr Phe Val Leu Arg Val Arg
705 710 715 720

Ala Gln Asp Pro Pro Pro Glu Leu Tyr Phe Val Lys Val Asp Val Thr
725 730 735

Gly Ala Tyr Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala
740 745 750

Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val
755 760 765

Val Gln Lys Ala Ala His Gly His Val Arg Lys Ala Phe Lys Ser His
770 775 780

Val Ser Thr Leu Thr Asp Leu Gln Pro Tyr Met Arg Gln Phe Val Ala
785 790 795 800

His Leu Gln Glu Thr Ser Pro Leu Arg Asp Ala Val Val Ile Glu Gln
805 810 815

Ser Ser Ser Leu Asn Glu Ala Ser Ser Gly Leu Phe Asp Val Phe Leu
820 825 830

Arg Phe Met Cys His His Ala Val Arg Ile Arg Gly Lys Ser Tyr Val
835 840 845

Gln Cys Gln Gly Ile Pro Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys
850 855 860

Ser Leu Cys Tyr Gly Asp Met Glu Asn Lys Leu Phe Ala Gly Ile Arg
865 870 875 880

Arg Asp Gly Leu Leu Leu Arg Leu Val Asp Asp Phe Leu Leu Val Thr
885 890 895

Pro His Leu Thr His Ala Lys Thr Phe Leu Arg Thr Leu Val Arg Gly
900 905 910

Val Pro Glu Tyr Gly Cys Val Val Asn Leu Arg Lys Thr Val Val Asn
915 920 925

Phe Pro Val Glu Asp Glu Ala Leu Gly Gly Thr Ala Phe Val Gln Met
930 935 940

Pro Ala His Gly Leu Phe Pro Trp Cys Gly Leu Leu Leu Asp Thr Arg
945 950 955 960

Thr Leu Glu Val Gln Ser Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile
965 970 975

Arg Ala Ser Leu Thr Phe Asn Arg Gly Phe Lys Ala Gly Arg Asn Met
980 985 990

Arg Arg Lys Leu Phe Gly Val Leu Arg Leu Lys Cys His Ser Leu Phe
995 1000 1005

Leu Asp Leu Gln Val Asn Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr
1010 1015 1020

Lys Ile Leu Leu Leu Gln Ala Tyr Arg Phe His Ala Cys Val Leu Gln
1025 1030 1035 1040

Leu Pro Phe His Gln Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg
1045 1050 1055

Val Ile Ser Asp Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys
1060 1065 1070

Asn Ala Gly Met Ser Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro
1075 1080 1085

Ser Glu Ala Val Gln Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu
1090 1095 1100

Thr Arg His Arg Val Thr Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr
1105 1110 1115 1120

Ala Gln Thr Gln Leu Ser Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala
1125 1130 1135

Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile
1140 1145 1150

Leu Asp

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGGATGATT CTTGTTGGTG ACAC

24

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AGCTTGGTAC CAACATGGAC TACAAGGACG ACGATG

36

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AATTCCCTTG TCATCGTCGT CCTTGTAGTC CATGTT

36

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGTTTGGTGG CTGATTCTT GTTGGTGAC

29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GTCACCAACA AGAAATCAGC CACCAAACG

29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GTCACCAACA AGAAAGCATC CACCAAACG

29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTCACCAACA AGAAAGCAGC CACCAAACG

29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GAATTCTAGA TCACTTGTCA TCGTCGTCCCT TGTAGTCGTC CAGGATGGTC TTGAAGTC

58

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CGTTTGGTGG CTGATTTCTT GTTGGTGAC

29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGTTTGGTGG ATGCTTTCTT GTTGGTGAC

29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CGTTTGGTGG CTGCTTCCTT GTTGGTGAC

29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ser Glu Ala Glu Val Arg Gln His Arg Glu Ala Arg Pro Ala Leu Leu
1 5 10 15

Thr Ser Arg Leu Arg Phe Ile Pro Lys Cys
20 25

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Arg Ser Lys Arg Arg Ser Arg Gln Pro Pro Arg Pro Gln Lys Thr Glu
1 5 10 15

Arg Pro Phe Ser Glu Arg Gly Lys
20

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Asp Pro Asp Ala Ser Gly Thr Phe Arg Ser Cys Pro Pro Glu Ala Leu
1 5 10 15
Lys Asp Leu

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CCCGGGTGGC GGAGGGTGGG C

21

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CGACTTTGGA GGTGCCTTCA

20

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGGAAGCTTT AATACGACTC ACTATAGGGT GGGCCTGGGA G

41

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCCGGGGGTTT CACAAAGCCCC C

21

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GGGAAGCTTT AATACGACTC ACTATAGGGG GTTCACAAGC CCCC

44

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Oligo nucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CCCGGGTGAG CCTGGGAG

18

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Arg Phe Ile Pro Lys
1 5

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Gly Ile Pro Gln Gly Ser
1 5

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Leu Leu Leu Arg Leu Val Asp Asp Phe Leu Leu
1 5 10